6-21-2013 12:00 AM

Semi-automatic Tracking of the Hyoid bone and the Epiglottis Movements in Digital Videofluoroscopic Images

Seereen Noorwali
The University of Western Ontario

Supervisor
Dr.Mahmoud El-Sakka
The University of Western Ontario

Graduate Program in Computer Science
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

Semi-automatic Tracking of the Hyoid bone and the Epiglottis Movements in Digital Videofluoroscopic Images

(Thesis format: Monograph)

by

Seereen Noorwali

Graduate Program in Computer Science

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada
June, 2013

Abstract

Swallowing is a process that happens hundreds of times per day during eating, drinking, or swallowing saliva. Dysphagia is an abnormality in any stage of the swallowing process. It can cause serious problems such as dehydration and respiratory infection. In order to help dysphasic patients, radiologists need to evaluate the patient's swallowing ability, usually using *Video Fluoroscopic Swallowing Study* (VFSS). During the assessment, several measurements are taken and evaluated, such as the displacement of the hyoid bone and epiglottis. Usually radiologists perform evaluation by means of visual inspection, which is a time consuming process that produces subjective results. Previous research has made strides automating swallowing measurements in order to produce objective results, but there is no study that automatically tracks the movement of the epiglottis. This thesis presents a design and implementation of a *Computer Aided Diagnosis* (CAD) system that can automatically track the movement of the hyoid bone and the epiglottis using minimal user input. The correlation between these two movements will be studied. With the aid of this system, radiologists can more reliably and efficiently take measurements and evaluate the health of the swallowing process.

Keywords: Dysphagia, Swallowing disorder, Videofluoroscopic Swallowing Study, Tracking, Hyoid bone, Epiglottis, SURF features, Template-Matching, Optical flow, correlate movements.

Acknowledgements

All praises belong to ALLAH, Alhamdulillah, who has given me the health, strength, and knowledge to finish this thesis successfully.

It is my pleasure to acknowledge King Abdullah bin Abdulaziz, the Saudi government, and Umm Al-Qura University for their financial support and supervision throughout my studies via the Saudi bureau in Canada.

I would like to express my deep gratitude to my supervisor, Dr. Mahmoud El-Sakka for his patient direction, enthusiastic encouragement, and useful suggestions, without which this thesis would not be complete. His academic and moral support during these past two years has been unceasing. I will forever be grateful to him. Thank you Dr. Mahmoud for your generous spirit. I am humbled and grateful that you consider me as your own daughter.

I wish to thank a few people for their support in helping me decipher the necessary medical details. Thank you to Professor Mandar Jog and Ms. Angela Roberts-South for providing me with the essential medical information and data that were used in this research.

I especially thank my father Dr. Mohammad Taher and my lovely mother Sabah for their care, their prayers, and their concern for my progress throughout my period of living abroad. I will forever be thankful to my grandmother and my sisters for their moral support and strong encouragement to complete this program. I want to thank my brother Hussain and his wife for their anatomical and medical resource support.

I would also like to extend a special thanks to my brother Abdulfattah and his wife; their support played a critical role in keeping me patient in this hard time. I greatly appreciate their inexhaustible kindness, patience, and love. I value their help, support, and encouragement over the past year. I will be forever grateful to them for welcoming me into their new family during my studies.

It is my pleasure to thank Ishtiaque Hossian for his help in the medical and technical information during the last two years. Also, I would like to thank Mark Dittmer for proofreading some of my thesis chapters.

Finally, I would like to thank my cousins and friends all over the world who have continued to support me morally and encourage me throughout my studies.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Swallowing is a process in the human body that regulates the movement of solids or liquids traveling from the mouth through the pharynx and esophagus to the stomach. Normally, people swallow hundreds of times a day while eating, drinking or swallowing saliva. Such regularity might lead one to think that the swallowing process is a simple one, but in fact this process is complex. It depends on multiple neural pathways, muscles and bones working in concert to make the traveling process successful.

Swallowing abnormalities (referred to as dysphagia) can cause swallowing difficulties. Failure to deliver the full volume of food via the normal route to the stomach is an indicator of dysphagia. Dysphagia can happen at any age, but is most common in seniors. There are some diseases that can cause dysphagia, such as Parkinson Disease, head and neck cancer and stroke.

In order to assist the patients who have swallowing difficulties, radiologists require a reliable method for evaluating patients' swallowing process. During the evaluation process, radiologists take several measurements that can be extracted from medical imaging. These measurements include timing the movement of food through the body, and timing and locating of different parts of the body during the swallowing process. Such measurements can be used as indicators to the health of various stages of the swallowing process. For example, maximum displacement of the hyoid bone and the rotation angle of the epiglottis through the swallowing process are important measures in the pharyngeal stage.

Radiologists use *Video Fluoroscopic Swallowing Study* (VFSS) to take sequential X-ray images of the head and neck area during the swallowing process. The patient is seated in front of a camera, and is asked to swallow different kinds of food mixed with barium (which is a contrast agent that is visible on X-ray images). By this means, radiologists can monitor the swallowing process and take measurements using the resulting series of X-ray images.

1.1. Inspiration

In the current state of affairs, radiologists typically take measurements via visual inspection. One problem with visual inspection is that it produces subjective results, leading to intra- and inter-rater variation problems. Furthermore, visual inspection is expensive, complex and time consuming.

In order to help radiologists produce objective results, several researchers have attempted to automate or semi-automate the extraction of measurements from swallowing footage. Some research uses computers in very simple ways, such as using the mouse to track the points in all frames, while other applications are more advanced. One gap in the research is that there is no study that automates the tracking of the epiglottis movement.

It is clear that the radiologists need a computer assistant to make their job easier and their results more reliable. One measure of ease of use is the amount and complexity of input required from the radiologist to allow the system to take measurements. Reliability can be measured by comparing multiple measurements (either by the same radiologist or different radiologists) from the same footage.

1.2. Contributions

This thesis demonstrates a design and implementation of a *Computer Aided Diagnosis* (CAD) system that can automatically track the movement of the hyoid bone and the epiglottis using minimal user input. The hyoid bone and the epiglottis play a significant role in the swallowing process. During the swallowing process, the hyoid bone moves upward and forward, opening a slit in the upper esophageal sphincter. The epiglottis is tilting down to cover the larynx, preventing food from entering the air channel. The correlation between these two movements is crucial for a healthy swallowing process. The CAD system presented here assists radiologists to study this correlation.

The CAD system performs a series of image processing steps to assist the radiologist. First of all, it identifies the vertebrae area in the X-ray image; this helps to identify the hyoid bone and epiglottis area. Tracking the vertebrae movement allows the system to measure according to axes consistent with the patient's location in the image. These axes allow the system to isolate the patient movement from the hyoid bone and epiglottis' movements. Once the vertebrae are identified, the system locates the hyoid bone and tip of the epiglottis so that they can be indicated to the radiologist in all frames. Finally, the system correlates the hyoid bone movement with the epiglottis movement to approximate the base of the epiglottis location in all frames.

Note that the proposed method is a CAD system that does not independently offer a diagnosis. Instead, it assists radiologists to take swallowing measurements.

1.3. Outline

The rest of the thesis contains four chapters. Chapter 2 describes the swallowing process and its assessment in detail. This chapter also describes several medical and engineering studies that deal with the swallowing process. Chapter 3 describes the proposed method with the experiments that have been done to adjust the parameters. It also describes the data set that is used in this research. Chapter 4 describes a series of experiments and results for the proposed CAD system. Finally, Chapter 5 draws conclusions from these experimental results, noted some limitations of the proposed method, and outlines some directions for future work.

Chapter 2

Background

2.1. Introduction

Swallowing is a complex process that involves various muscles movement and neural processes. It has several stages to allow the bolus[1] to travel from mouth to stomach. Any abnormality in any stage of the process can cause a swallowing disorder known as dysphagia, which can lead to severe problems. For the purpose of assisting dysphasic patients, a *Videofluoroscopic Swallowing Study* (VFSS) can be used to measure the patient's ability to swallow. VFSS involves specific measurement calculations that are related to the time and the movement of the food inside the body. The limiting factor in performing VFSS efficiently is visual inspection to analyze the videos that capture the swallowing process. Recently, several research groups have been developing new computer systems to assist radiologists to analyze these videos semi-automatically. In

[1] Bolus is a medical term for a soft mass of chewed food (http://www.merriam-webster.com/dictionary/bolus)

this chapter, the swallowing disorder and recent methods for assessing it are explained.

2.2. Medical Background

The brain is considered the central organ in the nervous system in humans. Controlling the operation of other biological components is one of the brain's fundamental functions; e.g., the brain controls the operation of the digestive system, the circulatory system, and the respiratory system. As such, any abnormality in the brain can cause severe problems in other parts of the body, for example, a movement disorder in various organs or limbs can be traced to problems in the brain [1].

One of the diseases that leads to movement disorders caused by an abnormality in the brain is *Parkinson's Disease* (PD). PD has the following four primary symptoms: *tremor* (trembling in the face, jaw, hands, arms, or legs), *rigidity* (stiffness of the limbs and trunk), *bradykinesia* (slowness of movement) and *postural instability* (impaired balance and coordination) [1]. When these symptoms become acute, many patients with PD have difficulty with day-to-day activities such as talking, walking and swallowing. In this research, we are concerned with processing data captured from PD patients that may experience swallowing difficulty.

2.2.1. Normal Swallowing Process

Swallowing is a complex process that involves different levels of brain control and sequential activation and deactivation of various muscle groups, including oral, pharyngeal, laryngeal, and esophageal muscles. This complex process can be divided into four phases, namely, oral preparatory, oral, pharyngeal, and esophageal [33].

Figure 2.1 shows some of the body parts that are involved in the swallowing process. The following subsections describe swallowing phases paraphrased from Logemann in [29].

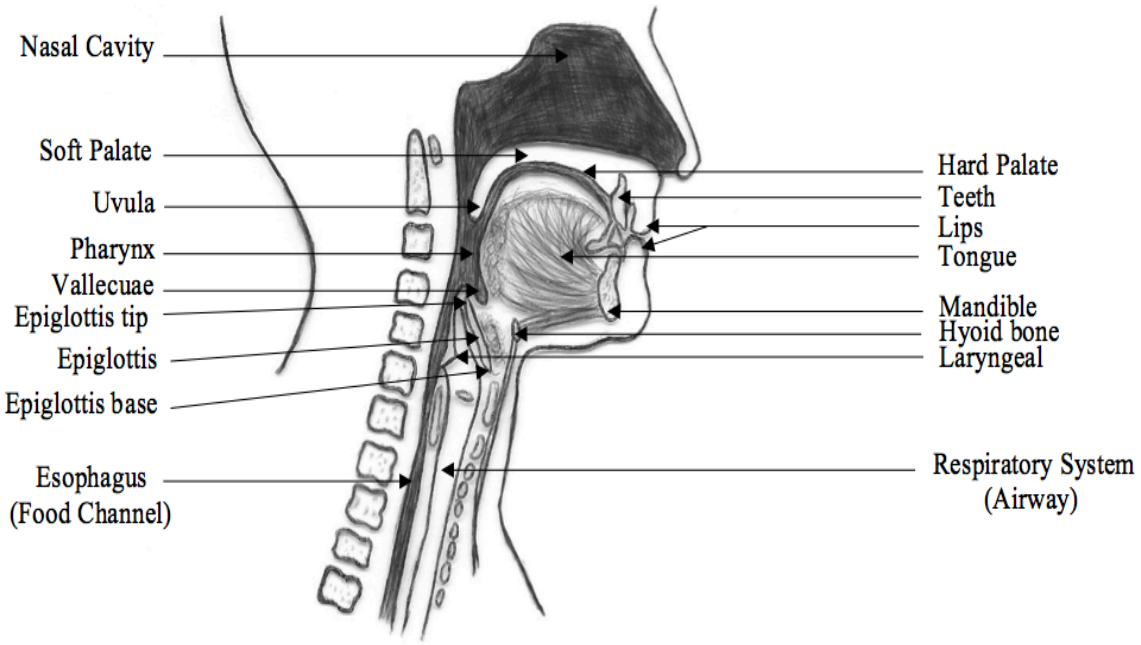


Figure 2.1: Some body parts involved in the swallowing process

2.2.1.1. Oral Preparatory phase

During the *oral preparatory phase* food is chewed and mixed with the saliva then formed into a ball called a bolus. The bolus is kept inside the mouth, as seen in Figure 2.2, under the hard palate between the front of the tongue, which is elevated to the lip and the back of the tongue which is elevated against the depressed soft palate.

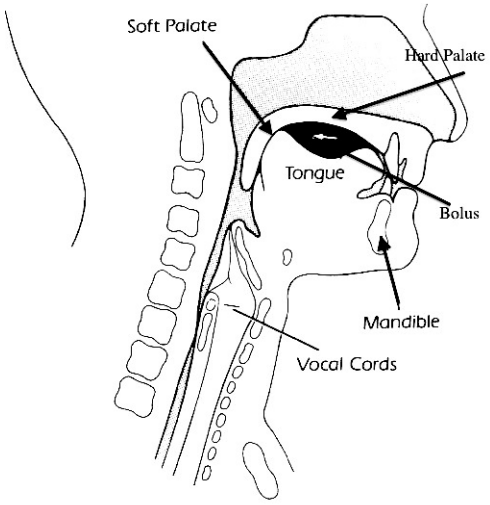


Figure 2.2: Oral Preparatory phase [29]

2.2.1.2. Oral phase

During the *oral phase* the tongue pushes the bolus to the back of the mouth via an anterior to posterior rolling motion (i.e., a front-to-back squeezing action). The back of the tongue is depressed as the front is elevated against the hard palate to push the bolus backward as shown in Figure 2.3. The *oral phase* ends when the bolus passes the anterior of the throat and touches the posterior wall of the pharynx.

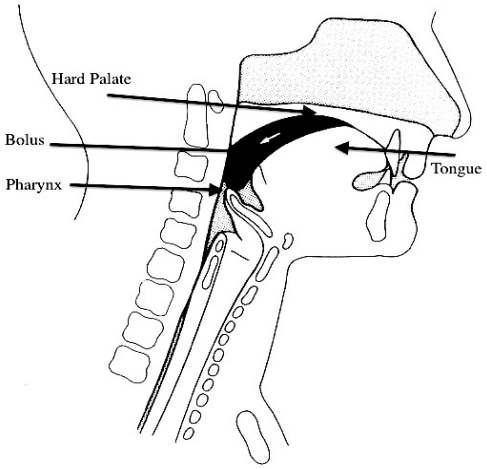


Figure 2.3: Oral phase [29]

2.2.1.3. Pharyngeal phase

During the *pharyngeal phase*, the airway is closed to prevent the entering of the bolus to the respiratory system. The soft palate is elevated and separated from the tongue to open a slit in the upper pharynx. At the same time, the nasal aperture is closed as a result of the elevation of the soft palate to prevent the bolus from entering into the nasopharynx region. In addition, the larynx is elevated and the laryngeal aperture is closed by the epiglottis to prevent the food from entering into the airway, as shown in Figure 2.4.

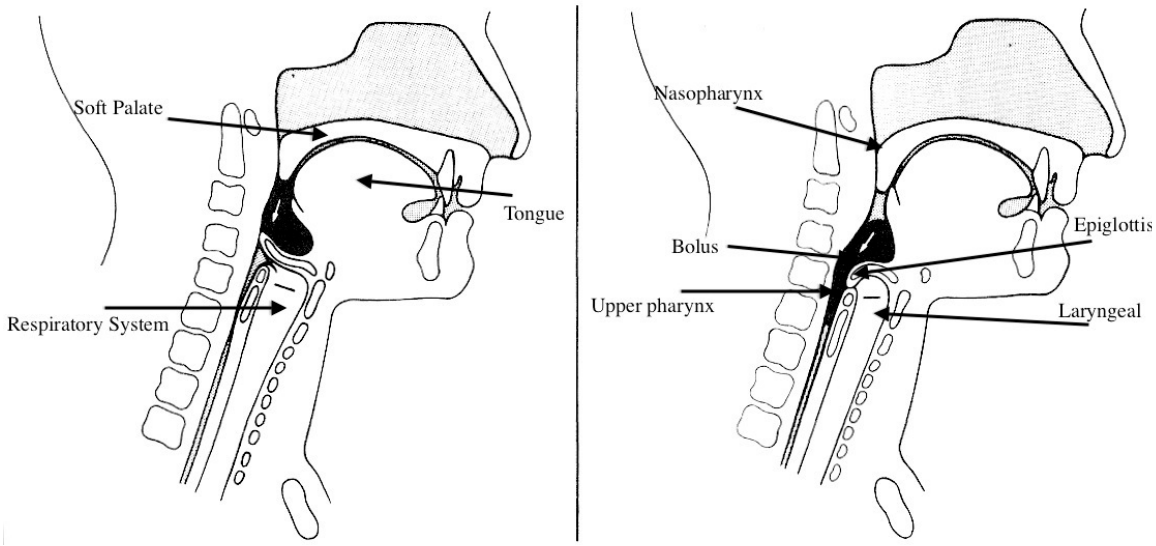


Figure 2.4: Pharyngeal phase [29]

2.2.1.3.1. Laryngeal Sub-phase

The *pharyngeal phase* contains a sub-phase called the *laryngeal sub-phase.* This sub-phase involves few concurrent actions. First, the hyoid bone and the larynx are pulled upward and forward to enlarge the pharynx, which creates a vacuum that pulls the bolus down. Second, the vocal folds adduct. Third, the epiglottis closes the top of the larynx to force the bolus to enter the esophagus.

2.2.1.4. Esophageal phase

Finally, during the *esophageal phase*, the bolus travels down the esophagus to the stomach, propelled by a squeezing action of the throat muscles, as shown in Figure 2.5 [29].

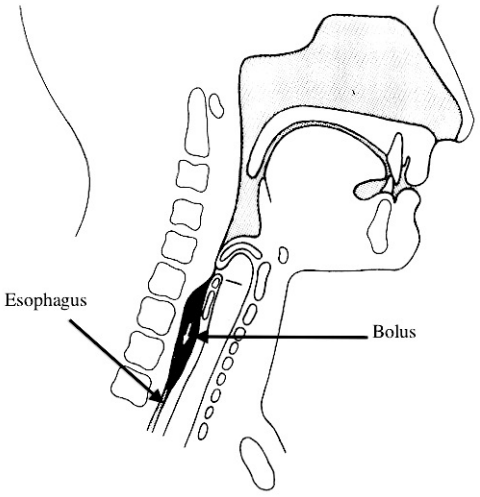


Figure 2.5: Esophageal phase [29]

2.2.2. Dysphagia

Dysphagia, difficulty in swallowing, manifests as one or more abnormalities in any phase of the swallowing process. It can cause serious problems such as malnutrition, dehydration, and respiratory infection. There are several conditions that can cause dysphagia, including stroke, head and neck cancer, multiple sclerosis, myasthenia gravis and PD.

2.2.3. Evaluation of Swallowing

To evaluate the ability of swallowing, there are several measurements that need to be calculated in each phase in the swallowing process. For example, in the pharyngeal

phase, tracking and analyzing of the hyoid bone's horizontal and vertical movements, as well as the duration of these movements are very important measures. These movements are called the *hyolaryngeal excursion* [33].

In addition, in the pharyngeal phase, tracking the epiglottis movement by calculating the rotation angle of the epiglottis and the speed of the rotation are important measurements to evaluate the swallowing process. Moreover, the movement of the hyoid bone is related to the movement of the epiglottis. Consequently, any abnormality in the hyoid bone movement can cause abnormality in the epiglottis movement [16],[18].

The *transit time* is another important class of measurements, which can be assesses by calculating the time that the bolus needs to pass specific regions such as oral and pharyngeal [8]. The *Pharyngeal Transit Time* (PTT) is defined as the time that the bolus takes to pass the region between the ramus of the mandible and the cricopharyngeal [3] (i.e., from the ramus of the mandible to the upper esophageal sphincter). The *Oral Transit Time* (OTT) defined as the time taken from the beginning of the backward movement of the bolus until the bolus tail passes the lower edge of the mandible (i.e., until it crosses the tongue base) [3].

In addition, it is important to measure the amount of bolus residue after swallowing in several regions, such as oral (*oral stasis*), pharyngeal (*pharyngeal stasis*), valleculea (*valleculea stasis*) and pyriform recesses (*pyriform recess stasis*) regions.

Aspiration and Penetration time are also important measures, which can be calculated as the time prior to, during and after the swallow when a bolus enters the airway.

Moreover, there are additional objective measurements that can be calculated from the movement of some soft tissues. These measures are mentioned in Baijens, *et al.,* [6] and summarized in Table 2.1.

Table 2.1: Several objective measurements with definitions [6]

Name	Definition
GPJo (*Glossopalatal junction opening*)	Moment of separation of the tongue and the soft palate
GPJc (*Glossopalatal junction closure*)	Moment of contact of the tongue and the palate after bolus propulsion
GPJd (*Glossopalatal junction duration*)	Δt between GPJo and GPJc
VPJc (*Velopharyngeal junction closure*)	Moment of the first contact of the soft palate against the posterior pharyngeal wall
VPJo (*Velopharyngeal junction opening*)	Moment of separation of the soft palate and the posterior pharyngeal wall with re-entry of air in the retrolingual space from the nasopharynx
VPJd (*Velopharyngeal junction duration*)	Δt between VPJc and VPJo
LVc (*Laryngeal vestibule closure*)	Moment when laryngeal elevation results in making contact between the arytenoid cartilages and the underside of the epiglottis
LVo (*Laryngeal vestibule opening*)	Moment of separation of the arytenoid cartilages and the underside of the epiglottis with re-entry of air in the laryngeal vestibule
LVd (*Laryngeal vestibule duration*)	Δt between LVc and LVo
UESo (*Upper esophageal sphincter opening*)	Moment of opening of the esophagus with entry of either air or barium
UESc (*Upper esophageal sphincter closure*)	Moment of closure of the esophagus after bolus transport
UESd (*Upper esophageal sphincter duration*)	ΔT between UESo and UESc
GPJo-LVc	ΔT between GPJo and LVc
GPJo-UESo	ΔT between GPJo and UESo
GPJo-UESc	ΔT between GPJo and UESc
Aspiration-penetration	Moment of aspiration or penetration

In addition, there are a group of subjective measurements mentioned in the same paper [6]. These measurements are evaluated using three-, five- or eight-point Likert scales,

where a smallest score refers to the normal ability of swallowing. These measures are summarized in Table 2.2.

Table 2.2: Several subjective measurements with definitions [6]

Name	Definition	Rating scale
Pre-swallow anterior spill	Preswallow loss of bolus from the lips	Five-point scale (0–4)
Pre-swallow posterior spill	Preswallow loss of bolus into the pharynx	Five-point scale (0–4)
Lingual pumping	Preswallow involuntary repetitive tongue movements	Five-point scale (0–4)
Swallow hesitancy	Delayed onset oral transport	Three-point scale (0–2)
Piecemeal deglutition	Sequential swallowing on the same bolus	Five-point scale (0–4)
Delayed initiation pharyngeal reflex	Delayed onset pharyngeal triggering	Three-point scale (0–2)
Post-swallow oral residue	Postswallow pooling in the oral cavity	Five-point scale (0–4)
Post-swallow vallecular pooling	Postswallow pooling in the valleculae	Three-point scale (0–2)
Post-swallow pyriform sinus pooling	Postswallow pooling in the pyriform sinuses	Three-point scale (0–2)
Penetration aspiration scale (PAS)	Penetration and/or aspiration	Eight-point scale (1–8)

2.2.4. Clinical Assessment of the Dysphagia

During the clinical swallowing assessment process, several modalities can be utilized to evaluate the ability of swallowing, e.g., ultrasonography and videofluoroscopy. In Yabunaka, *et al.* [44], ultrasonography is used to evaluate the hyoid bone movement because it is inexpensive and has no known side effects. Due to the large amount of

noise involved in the produced ultrasound images, the visual inspection of videofluoroscopy frames is still considered the *gold standard* method.

VideoFluoroscopy or VideoFluoroscopic Swallowing Study (VFSS) is currently considered the most precise and accurate method for the capturing the swallowing process for clinical assessment. It is defined as dynamic X-ray images of swallowing that combine traditional fluoroscopy with video technology to capture the motion and record it. Using this video, the swallowing process is reviewed by a specialist in real time, frame-by-frame, or in slow motion. To capture a fluoroscopic video of the swallowing process, the subject sits upright in front of an X-ray machine while the camera records a lateral view of the head and neck area, as shown in Figure 2.6. During this recording process, subjects are asked to eat food mixed with barium to make the bolus visible in the X-ray video. The radiologist usually gives the subject soft food, hard food or liquids in various quantities. The X-ray machine is turned on only during the swallow to limit the radiation dose. See Palmer *et al.* [34] for a detailed description of the VFSS procedure.

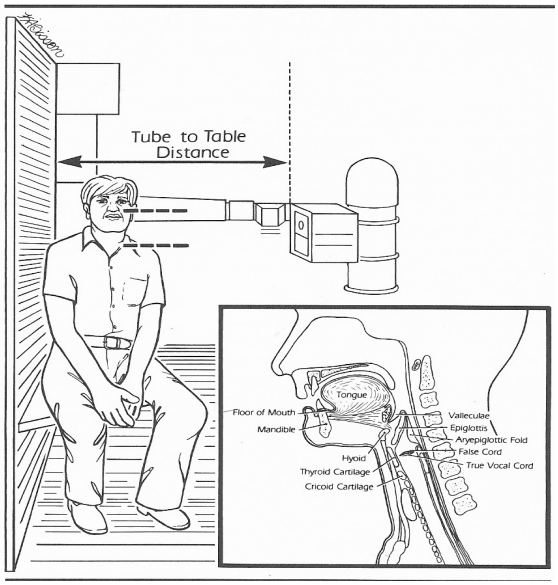


Figure 2.6: Patient seated on the platform attached to the fluoroscopy table so that the upper aerodigestive tract can be viewed laterally [29]

2.2.5. Medical Studies Background

Currently, radiologists still perform visual inspection to assess and examine videofluoroscopic images. Many medical studies rely on radiologists' visual inspection to calculate some of the measurements like hyoid bone movement, transit time and bolus residual in several regions.

There are many studies that assess various measures that indicate a swallowing disorder through visual inspection of videofluoroscopic images. Rosenbek *et al*. [39] evaluate the accuracy of an eight-point Likert scale to indicate the penetration and aspiration during the swallowing process. Kahrilas *et al.* [23] analyze the mechanism of volume accommodation. Robbins *et al.* [36] use the same scale in [39] to evaluate airway protection. Eisenhuber *et al.* [15] evaluate the clinical significance of pharyngeal retention to predict aspiration in patients with dysphagia. Moreover, Kendall *et al.* [25] identified abnormalities of the timing and extent of *Upper*

Esophageal Sphincter (UES) opening in dysphasic elderly patients. Perlman *et al.* [35], determine the relationship between events observed with simultaneous videofluoroscopy and respirodeglutometry. In addition, Gokyigit *et al.* [17] investigate the time interval between glottis closure and the opening of upper esophageal sphincter. Rofes *et al.* [37] diagnose oropharyngeal dysphagia in the elderly. Choi *et al.* [10] investigate the mechanisms of aspiration with respect to the viscosity of ingested material.

Some of medical studies focus on dysphasic patients under specific medical conditions, For example, several studies focus on dysphasic patients with stroke [19],[11],[43]. Kluin *et al.* [26] evaluate the dysphagia in elderly with myasthenia gravis. In addition, Eisbruch *et al.* [14] study patients with head and neck cancer. Sun *et al.* [41] study patients with globus pharyngeus. Higo *et al.* [20] assess patients with myasthenia gravis. Jang *et al.* [22] measure the ability of swallowing for patients with vocal cord paralysis. Namaki *et al.* [32] focus on patients with oral cancer. Leonard *et al*. [28] evaluate swallowing parameters in dysphasic patients after anterior cervical-spine surgery.

Several studies attempt to measure the reliability of analysis the videofluoroscopic images. Kuhlemeier *et al.* [27] measure the intra- and inter-rater variation in the evaluation of VF swallowing. Mccullough *et al.* [31] measure the inter- and intra-judge reliability of videofluoroscopic footage. Finally, Stoeckli *et al.* [40] assess the interobserver reliability of videofluoroscopy for swallow evaluation.

Additionally, some studies evaluate the accuracy of computational methods of

obtaining relevant measures from videofluoroscopic footage from a medical perspective. Rommel *et al.* [38] studied the growth of the oropharynx and hypopharynx in infants and young children using a computer program designed specifically for this purpose. Also, Dyer *et al*. [13] calculate the bolus residue in the valleculae through computer assessment.

It is important to note that even though several of the above mentioned studies apply some level of computer assessment, all of them involve visual inspection and manual assessment. The main drawback of visual inspection of videofluoroscopic images is its subjective nature; radiologists usually face intra- and inter-observer repeatability issues and bias when taking measurements [4]. Moreover, the visual inspection process is expensive, complex, and time consuming. Consequently, there are several research programs aimed at automating the evaluation of the swallowing process.

2.3. Computer Aided Diagnosis (CAD) Background

Computer Aided Diagnosis (CAD) is an approach to engineering diagnostic tools, which assumes that there is a role for both human experts and computers in an efficient and accurate diagnostic process. CAD is different from automatic computer diagnosis, which attempts to replace human diagnosticians completely [12]. In most CAD systems, digital medical images are analyzed by computer systems; however, the result is used as a second opinion to assist physicians to make their final diagnostic decisions. In the next subsections, recent CAD systems for analyzing the swallowing process are summarized.

2.3.1. Calculating OTT and PTT

Aung *et al.* [3] introduce new software that helps radiologists to quantify OTT and PTT (described in Section 2.2.3) using two boundary lines or landmarks that separate the oral cavity from the pharynx and the pharynx from the esophagus, respectively, as shown in Figure 2.7.

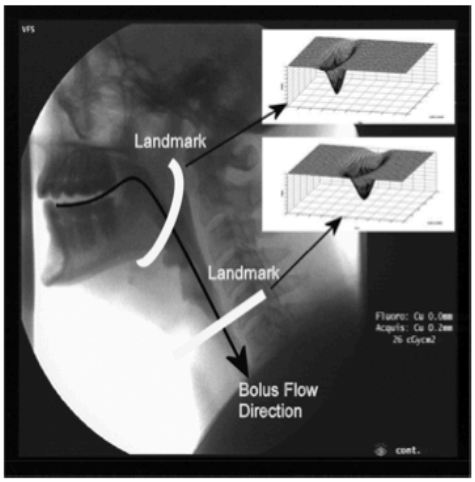


Figure 2.7: The first step in this system is to initialize the landmarks at the showing locations [3]

After initializing the landmarks by user interaction, the registration process is applied to update the coordinate of the landmarks in each frame in order to compensate for the patient's movement during the swallowing process. Then, a spatiotemporal plot can be generated. The intensity and the rate of change of intensity of the pixels at the landmark are plotted against both time (frame number) and location along the length of the landmark. When the bolus passes by the landmark, a black area appears on the spatiotemporal plot.

Finally, the system is automatically determining the frame number, α, at which the

bolus head arrives at the landmark and the frame number β, at which the bolus tail passes the landmark. These frame numbers can be used to calculate PTT and OTT. The correlation coefficients between the results obtained by the three oral pathologists and the CAD system are calculated.

A second study [5] implements the same idea as in the previous study [3]; however, it has some alternative methods applied in the last stage (the way of calculating α and β frames). All beginning stages are the same as in the previous study [3].

2.3.2. Determining several boundaries automatically

Another study by Aung *et al.* [4] introduces a CAD system with minimal user input that could determine accurate locations for several boundaries of the anatomical swallowing region using a deformable shape template. This template can be automatically fit to each image using the *Active Shape Model* (ASM).

During the swallowing assessment, radiologists focus on several anatomical regions including the pharyngeal, laryngeal and hyoid bone. Therefore, a simple shape model is designed by marking the boundaries of these regions using only 16 points (eight straight lines). These points represent the vertebrae, the hyoid bone and the laryngeal walls as shows in Figure 2.8.

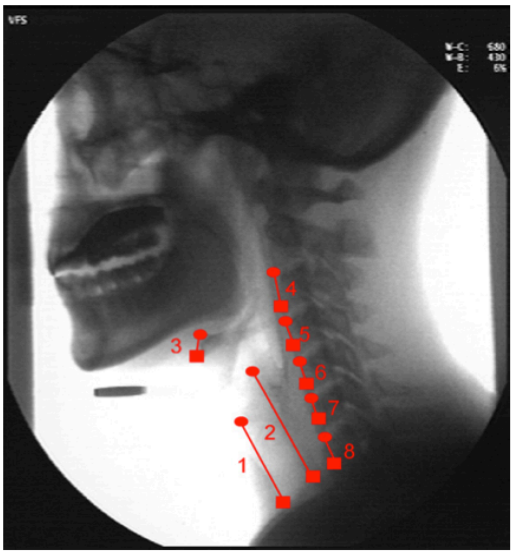


Figure 2.8: The 16-points active shape model demarcating 8 salient edges [4]

The training set of shapes is manually created by locating these 16 anatomical points in each image in a training database. To determine the statistics of the relative variations of shapes, the *Principal Component Analysis* (PCA) method is applied to find the principal axes of the ellipsoidal clusters.

Once the mean shape, range and modes of allowable shape variations are determined from the training set, the next step is to find the best fitting model in a new image. Overall, fitting was found to be more reliable on the vertebrae and inferior points of the larynx compared to the superior laryngeal points and hyoid bone.

2.3.3. Calculating the PTT, OTT and hyoid bone trajectory

Ceccarelli *et al.* [8] attempts to measure the swallowing parameters using a semi-automatic system. The main goal of this study is to calculate the OTT, PTT, and the trajectory of the hyoid bone movement.

For calculating the transit time, a radiologist defines two straight lines as the landmarks in the first frame, as shown in Figure 2.9. Then, the landmarks in each frame were calculated. Finally, PTT and OTT can be calculated after extracting the frame numbers at which the bolus arrives and passes the landmarks respectively.

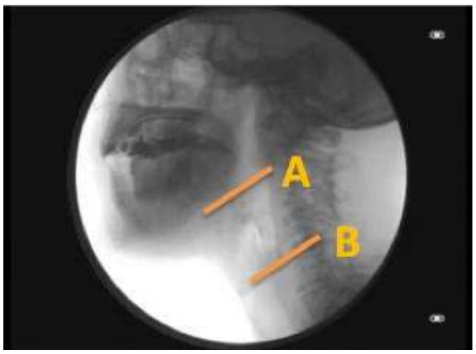


Figure 2.9: Landmarks positions in this system [8]

The second task in this system is tracking the hyoid bone movement. A radiologist chooses two frames that show the resting position for the hyoid bone (R-frame) and maximum displacement of the hyoid bone (M-frame). The radiologist should select 6 points in these frames to assist the system in calculating the horizontal and vertical displacement of the hyoid bone.

2.3.4. Oral movement

Chen *et al.* [9] proposed a CAD system that computes measurements of oral movement in videofluoroscopic images. The system can automatically track small metal markers attached to the subject's tongue and teeth by utilizing template matching. Generic templates of markers are designed by inspecting the shape, size and gray-value pattern of manually identified markers in the sample shape. The generic template is matched in the frames to find the location of the markers. Due to the head movement, the

movement markers are expressed with respect to a number of reference markers.

2.4. More medical background most related to this thesis

For our research, we have designed a CAD system that is focusing on tracking the hyoid bone and epiglottis movement in PD patients. The main goal of this research is minimizing the required input from the user (user interaction) to streamline radiologists' workflow. As described in Section 2.2.3, calculating the movement of the hyoid bone and epiglottis are important measures to evaluate the swallowing ability.

It is worth mentioning that, the hyoid bone is not connected with other bones; instead, it is attached to a group of muscles under the tongue. The shape of this bone is like a horseshoe (similar to letter the U). The surrounding muscles direct the hyoid bone during swallowing; this movement has both vertical and horizontal components. The epiglottis is a small flap of cartilage located at the base of the tongue under the mandible. The main function of the epiglottis is protecting the larynx by closing the air channel to prevent the bolus from entering the airway. In the next subsections, additional information about the hyoid bone and epiglottis that is pertinent to our research is reported.

2.4.1. Epiglottis

2.4.1.1. Movement of the epiglottis

The epiglottis can be seen in x-ray images as a leaf-shaped flap of cartilage tissue and lies just behind the root of the tongue [18]. It is on the right side of the hyoid bone and

the left side of the C2 - C3 - C4 cervical vertebrae, under the mandible and upper airway channel. The epiglottis can be seen in the white box in Figure 2.10.

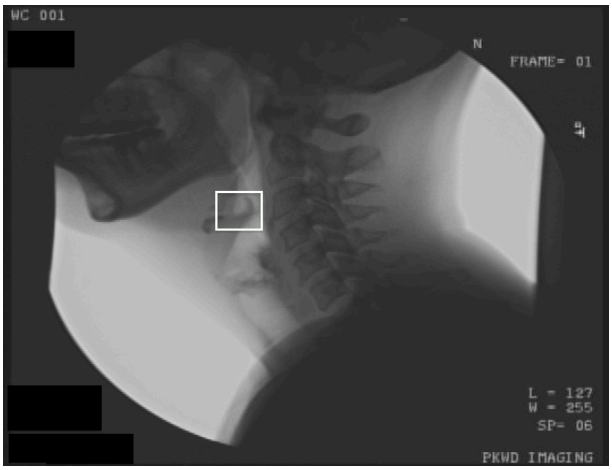


Figure 2.10: The epiglottis place in the x-ray (frame from the available data)

In normal swallowing, the epiglottis moves from its vertical resting position, tilts completely downward, and returns to its resting position in less than a second. The epiglottis tilts down in two steps during swallowing: The first movement is from an upright to a transverse position. The second movement is from a transverse position to inverted position when the epiglottis tilted completely into esophageal [16]. The tilting process is illustrated in Figure 2.11.

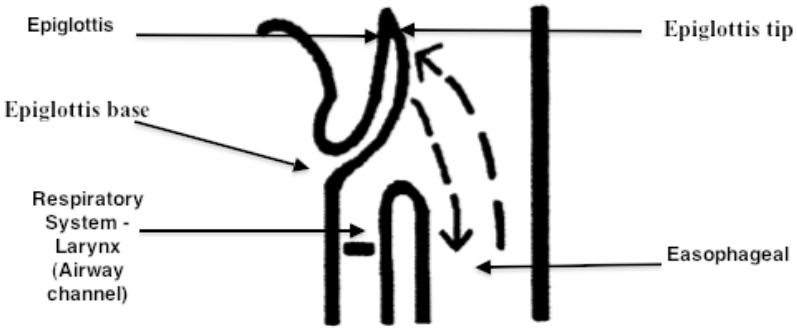


Figure 2.11: Normal movement of the epiglottis The epiglottis is in the resting position. The arrows show its movements. First, tilts completely down. Second, return back to resting position [18]

2.4.1.2. Epiglottis measurements

There are several studies that focus on epiglottis movement. Ekberg *et al.* [16] describe the normal epiglottis movement from anatomical point of view. In contrast, Garon *et al.* [18] identify the abnormal pattern for the epiglottis movement in the dysphasic patient. Pike *et al.* [33] calculate the angle of the epiglottis rotation for the normal subject and dysphasic patients from different epilogists. Their results indicate that the rotation angle of the epiglottis for the patient groups is significantly different than that of the normal group. See Figure 2.12 that shows the mean of the epiglottis rotation for normal and patient groups.

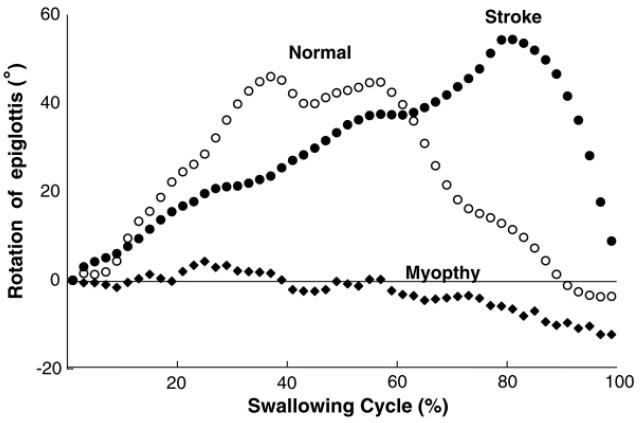


Figure 2.12: Mean for the rotation angle of the epiglottis [33]

Analyzing the epiglottis movement can yield several important measurements. First, the rotation angle can be calculated. The initial angle is 0 when the epiglottis in the resting position, the angles are calculated in the clockwise direction. To achieve this, a line from the base to the tip of the epiglottis is identified. Then, the rotation of the angle is computed. In addition, the maximum angle of the rotation can be estimated

when the epiglottis is completely downward (tilted). Finally, the speed of the rotation can be measured as well.

These measurements can be used to evaluate the epiglottis's ability to protect the airway and can help to measure the movement status (normal or not). In addition, identifying the movement pattern can help the radiologist to recognize the etiology of the dysphagia. Moreover, the speed of the rotation can be significant for determining aspiration status. The speed of rotation is estimated by measuring the time elapsed to complete one rotation (normal time for one rotation < 1 second) [16],[18].

2.4.2. Hyoid bone

2.4.2.1. Movement of the hyoid bone

In the X-ray, the hyoid bone can be seen as a U-shaped form just behind the epiglottis. The hyoid bone is under the mandible and in the left side of the epiglottis, which is on the left side of the C2 - C3 - C4 vertebrae. The hyoid bone can be seen in in the white box in Figure 2.13.

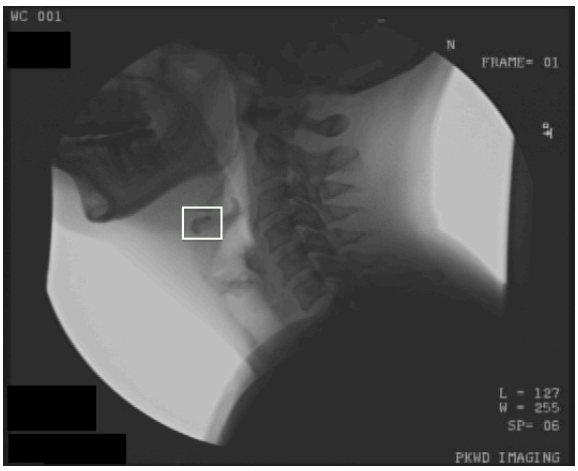


Figure 2.13: The hyoid bone place in the x-ray show in white box (available data)

The movement of the hyoid bone during swallowing has both vertical and horizontal components. During normal swallowing, the hyoid bone moves upward and forward, and then pauses in place for a short time. Then, the hyoid bone moves downward and backward to its original position. Figure 2.14 shows the normal trajectory of the hyoid bone.

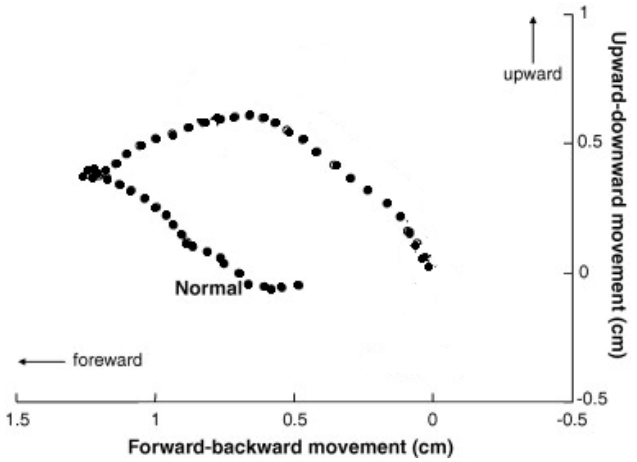


Figure 2.14: Pattern of normal movement of the hyoid bone (edit from [33])

2.4.2.2. Hyoid bone measurements

Tracking hyoid bone movement can yields several important measurements; horizontal and vertical displacement, as well as duration of movement indicate normality (or abnormality) of the swallowing process. In addition, speed of the horizontal and vertical displacement is essential to evaluate the hyoid bone ability during swallowing.

For tracking the hyoid bone, several studies have tried to track the hyoid bone movement manually. Paik *et al.* [33] examine the hyoid bone and epiglottis trajectories, comparing healthy subjects in a control group against dysphasic patients in an experimental group. In addition, Tsushima *et al.* [42] examine hyoid bone

movement during sequential swallowing and evaluate the relationship among trajectory patterns.

However, there are other studies that tried to track the hyoid bone movement semi-automatically by designing a special CAD system for that. Kellen *et al.* [24] focuses on the hyoid bone movement during the swallowing process. In their system, the radiologists should identify the hyoid bone on one frame, which is the *Region Of Interest* (ROI) as shown in Figure 2.15.

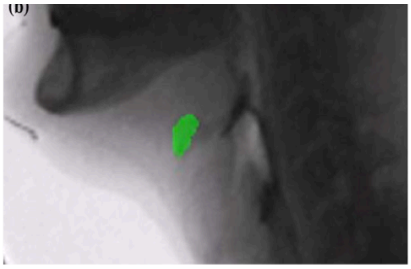


Figure 2.15: User defined template for the ROI (green pixels) approximately cover the hyoid bone and surrounding pixels [24]

The next step is tracking the template for the entire image sequence by matching the edges. To accurately track the hyoid bone with respect to the patient, Kellen *et al.* compute a patient-centric coordinate system. This system accounts for the subject's head and neck movement.

Hossain [21] designed a special semi-automatic CAD system that focused on tracking the hyoid bone movement from videofluoroscopic images. At the beginning of the process, the system identifies the beginning of the swallow segment automatically. Then, the ROI that has the cervical vertebrae was identified in all frames automatically by a classification-based approach that needs a special training. To identify the three cervical vertebrae and the hyoid bone that are located in the ROI for all frames,

template matching was utilizing after indicating the objects of interest by user input. The tracking coordinates for the hyoid bone were transformed to the other reference system associated with the patient's structure to isolate the patient's movement from the hyoid bone movement. The system produces successful results for tracking the hyoid bone.

2.4.3. Correlating hyoid bone and epiglottis together

There are several interconnected muscles and ligaments in the area between the epiglottis and the hyoid bone. The epiglottis is suspended between the adjacent structures by fibroelastic elements. The anterior aspect of the epiglottis is connected by the hyoepiglottic to the hyoid bone. This suspension of the epiglottis implies that the movement of the thyroid cartilage and the hyoid bone are promptly transmitted to the epiglottis [16].

An approach of this structure tilts the epiglottis downward to a transverse position, while structure subsequent separation returns the epiglottis to its resting and upright position. Therefore, the first movement of the epiglottis from upright to transverse position is accomplished by elevation of the hyoid bone and the thyroid cartilage. This movement is a passive one, induced by the muscles that lift the hyoid bone. The epiglottis's tilt downward (by 90 degree) is synchronized with the elevation of the larynx and hyoid bone [16].

The second movement of the epiglottis is from the transverse position to a position with its tip in the esophageal. This movement is the result of muscles acting directly on the epiglottis. This movement occurs later in swallowing and is related to contraction

of the thyroepiglottic muscle. During this phase of movement the epiglottis is tilled further down 30-90 degrees [16].

The available evidence shows that, the hyoid bone moves horizontally and vertically in sync with the downward-to-transverse movement of the epiglottis. The tip of the epiglottis tills down (to the transverse position) while the hyoid bone moves upward and forward[2]. Figure 2.16 shows the positions for the epiglottis and the two movements of the epiglottis.

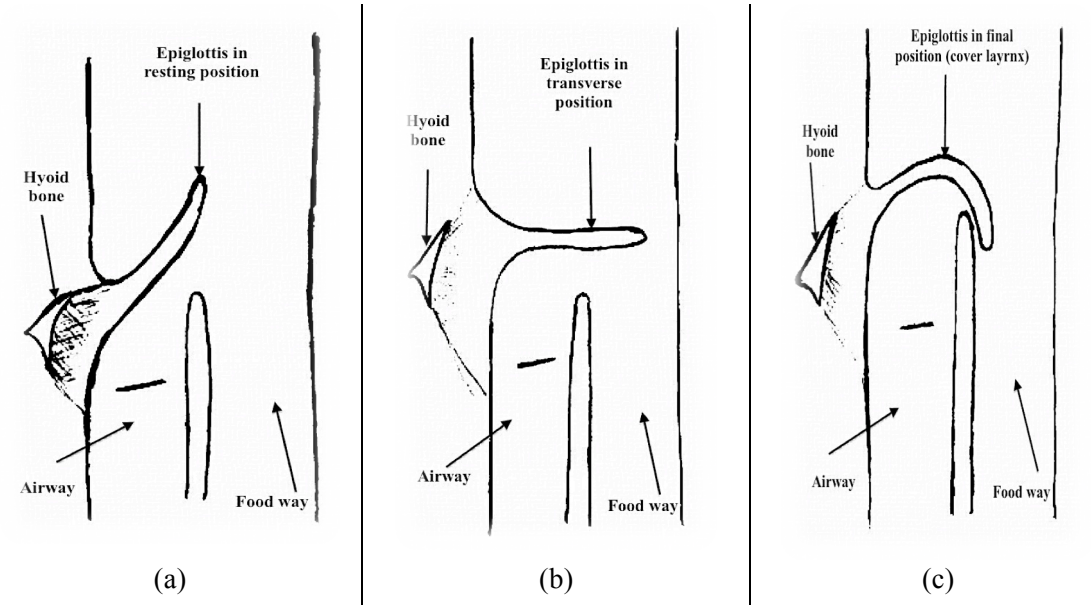


Figure 2.16: Three positions for the epiglottis. (a) Resting position, (b) Transverse position (c) Final position

[2]Note that the maximal position of the epiglottis is reached following both the downward-to-transverse movement and the further movement resulting from internal muscular contraction.

2.5. Summary

This chapter describes how to evaluate and assess a swallowing disorder (dysphagia) that can occur due to various diseases. Additionally, it discusses recent CAD studies for analyzing the swallowing process using VF. Although several groups work in this field, a complete automatic system has not been derived yet. Furthermore, several measurements to evaluate the swallowing process have not even been semi-automated. At the end of this chapter, we describe hyoid bone and epiglottis movement; automating measurements associated with this process is the focus of our research. Even though some studies tried to automate hyoid bone, there is no study that tried to automate the epiglottis movement. Furthermore, no one has tried to correlate the hyoid bone movement with epiglottis movement automatically.

Chapter 3

Data and Proposed Method

This chapter contains three significant parts. Section 3.1 describes the data that are used in this research in detail. In Section 3.2 the pre-processing steps are described. Finally, the actual processing is explained in Section 3.3.

3.1. Data

3.1.1. Data Acquisition

The data used in this research were acquired and stored using the *Picture Archiving and Communication System* (PACS). PACS is a medical imaging technology used by radiologists to capture, store and access image data from multiple sources.

In this research, the only modality we consider is X-ray images. After images are captured from an X-ray machine, the image data are archived as video sequences. These sequences are then recorded on a DVD disc which can be played in a typical

DVD player. Some of the characteristics of the videos recorded on the DVD discs are shown in Table 3.1.

Table 3.1: Characteristics of the videos in the original DVDs

Encoding scheme	MPEG-2
Frame format	YUV420P
Frame rate	30 fps
Frame size	720x480

3.1.2. Data Description

3.1.2.1. Data Format

The DVD media contains a number of files that serve different purposes. This study is only concerned with the files that contain image data; that is, only the files that end with a .VOB (Video OBject) extension. Each swallowing session (i.e., each series of swallowing video captures associated with a single patient) is divided into a number of VOB files to be recorded on the DVD disc. Each disc may contain more than one swallowing session. Each DVD disc contains several .VOB files. The names of these files started by the three letters VTS (*Video Title Set*) and follow the naming convention VTS_nn_*x*.VOB, where nn is from 1 to 99 and *x* is from 1 to 9. All VOB files are encoded in standard MPEG-2 video format.

3.1.2.2. Interface to the data

Our research software uses the FFMPEG[3] project for input and output operations on a DVD VOB file. FFMPEG is an open source software that provides facilities for handling multimedia data including video. FFMPEG offers a variety of features. The features used in our research are listed as follows:

- VOB encoding/decoding.
- Reading/writing from/to various types of video files formats (e.g., avi, vob)
- Accessing to various attributes of a video file (e.g., pixel format, frame numbers)
- Converting between various types of multimedia data (e.g., avi, pgm, jpeg)

In addition, in this research the OpenCV library is used. OpenCV (*Open Source Computer Vision*) is a library of programming functions for real time computer vision [2]. See Appendix B for a list of the main OpenCV functions that have been utilized in this research.

3.2. Data pre-processing

After the data are acquired, the videos in the DVD discs go through several pre-processing steps. Ishtiaque Hossain wrote software to perform the first pre-processing step required for this research project in 2012 [21]. This first step consists of removing patient information from frames. Hossain identifies frames containing patient

[3] The formal website for FFMPEG http://www.ffmpeg.org/

information using frame variance and replaces these frames with a special frame that is entirely black except for the first four white pixels (starting from the top-left of the frame). At the end of this step, he generated a folder that contains several AVI files that have been encoded using FFV1 lossless intra-frame video codec.

3.2.1. Removing black frames

Picking up where Mr. Hossain left off, the second pre-processing step is to remove the artificial black frames that have been added by Mr. Hossain's pre-processing step. These frames are easy to remove as they have fixed structure (four white pixels followed by black pixels). These mostly black frames have been removed to speed the process.

3.2.2. Generating single swallow videos

The third pre-processing step is to divide videos into several "single-swallow" segments, each one depicting exactly one swallow process. The purpose of this step is to simplify the coming processing steps for independent "single-swallow video". As explained in Section 2.2.4, each video is an X-ray of a human subject's head and neck, as the subject swallows different types of food that have been mixed with barium. The original data from the DVD contains continuous footage of multiple clips, each depicting a single swallowing processes. To make matters worse, sometimes a single swallowing process spans two AVI files, when a swallowing process beginning near the end of one file and ending near the beginning of the next. See Figure 3.1.

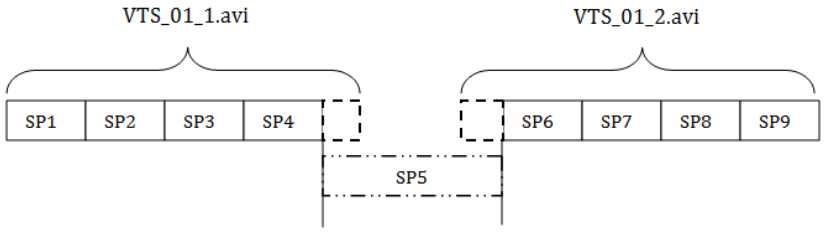


Figure 3.1: Example of the swallowing process order in a two avi files

In this pre-processing step, a single-swallow video file is created for each swallowing process depicted in the original footage; these video files have the technical characteristics as described in Table 3.2. The output of this pre-processing step is a series of single-swallow videos.

Table 3.2: Characteristic of the videos that generated in the third pre-processing

Encoding scheme	FFV1
Frame format	Gray8
Frame rate	25 fps
Frame size	720x480

This pre-processing step starts by identifying the beginning of the swallow. At a predefined location in the top-right corner of each frame, the X-ray machine adds a special marker to identify the frame as a beginning of a swallowing process or not. See Figure 3.2.

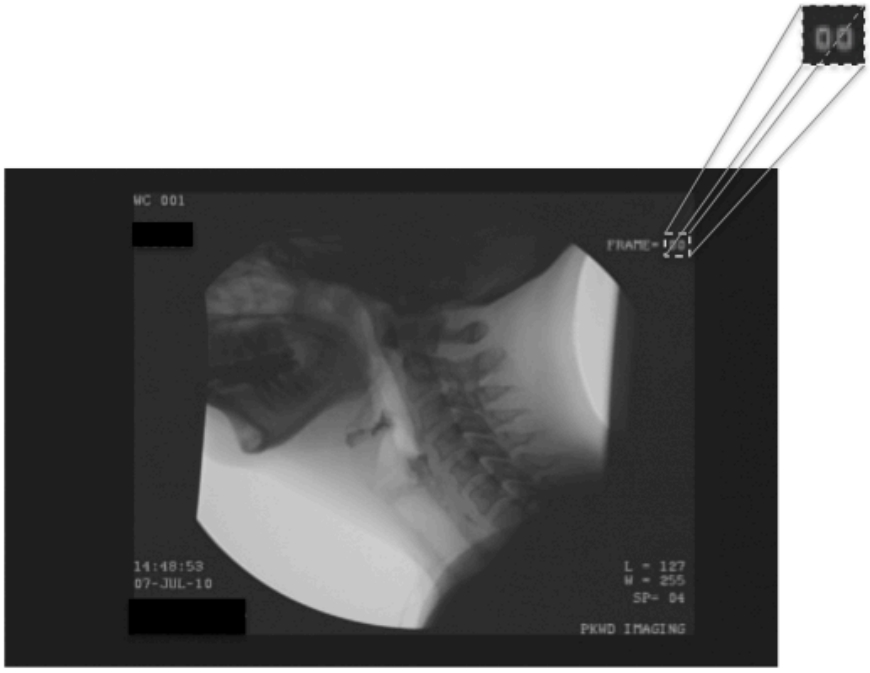


Figure 3.2: Marker location in a frame

For the first few frames of every captured swallowing process, the special marker is set to be a white two-digit number on a black background—we will refer to these frames and these markers as *number-frames* and *number-markers*, respectively (see Figure 3.3). The used marker in the rest of the frames of the swallowing process is a black "XX" on a light grey background—we will refer to these frames and markers as *XX-frames* and *XX-markers*, respectively (see Figure 3.4). As such, identifying the beginning of a swallow process involves analyzing the top-right corner of each frame to locate and recognize the beginning of the swallowing process in the video footage.

(a) (b)

Figure 3.3: Two examples of a number-marker

Figure 3.4: An Example of XX-marker

There are many methods that can be used to identify *number-frames*. Two of them are using template matching or calculate the average. The first method utilizes the correlation coefficient between the marker area of each frame and a template from a "typical" *XX-marker*, where a threshold is empirically defined and used to identify the frame type.

The second method that can be used to identify the *number-frames* is to utilize the average pixel value of the marker area. As shown in Figure 3.3 and Figure 3.4, the *XX-marker* is brighter than that in the *number-marker*. To demonstrate this observation, the averages of the pixel values in the marker areas for 20 random *XX-markers* and 20 random *number-markers* (see Figure 3.5) are calculated and plotted (see Figure 3.6).

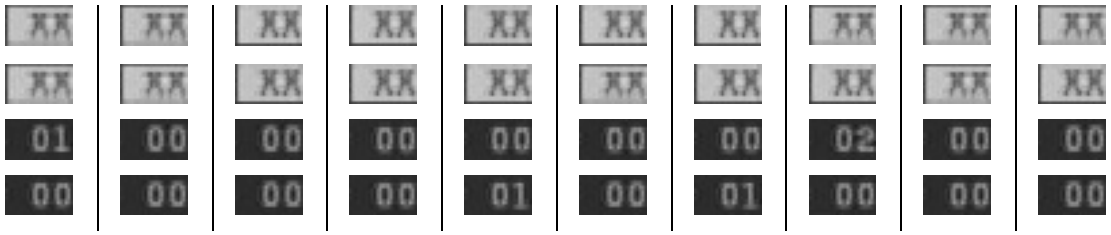


Figure 3.5: Samples of *xx-markers* and *number-markers* areas

The results illustrate that the averages of the *XX-marker* intensities do not exceed 163 and do not go less than 158. For the *number-marker*, the averages do not exceed 63 and do not go less 59. The threshold we use to differentiate between markers of each type is the "middle value" between 63 and 158 i.e., 110.5. That is, if the average intensity value of the marker area is less than 111, the frame is a *number-frame*. Otherwise, the frame is an *XX-frame*.

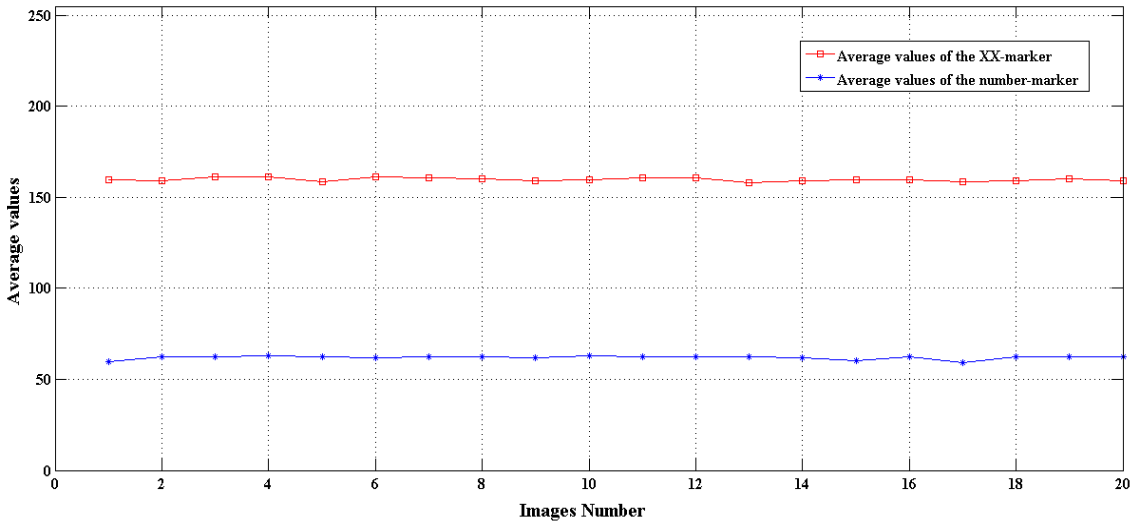


Figure 3.6: Average pixel values of the marker area in 20 random *XX-frames* and 20 random *number-frames*

The first method (correlation coefficient method) needs to load a template marker and do multiple operations on both images (input frames & template marker) each time to calculate the correlation coefficient. On the other hand, the second method (the average of the marker area method) is a simple procedure and does not require loading template marker for comparison. For the purpose of this thesis, the second method was applied, as it is more efficient and equally effective.

3.3. Processing Data

3.3.1. Locating main area for processing

The main area, or main region-of-interest, during all processing steps is referred as the *main-roi*. The *main-roi* is a center area that is about 45% of the original image size. The *main-roi* is framed by the text at the top and bottom of the image, and the black areas on left and right sides; that is to say, it is defined as the area below the text at the top, above the text at the bottom, and between the black areas to the left and right.

After doing simple experiments to identify the *main-roi* in individual patients images, it was determined that the top-right corner of the *main-roi* is situated at 120 pixels from the right edge and 80 pixels from the top edge, with a width of 525 pixels and a height of 300 pixels. Figure 3.7 shows the *main-roi* in one frame. The *main-roi* is fixed for all swallowing segments in all patients. Defining the *main-roi* as the region of interest for all processing steps reduces the search space for locating smaller points and areas of interest. Two functions from the OpenCV library are used to set and reset the *Region Of Interest* (ROI) in the images, cvSetImageROI and cvResetImageROI [2]. See Appendix B.

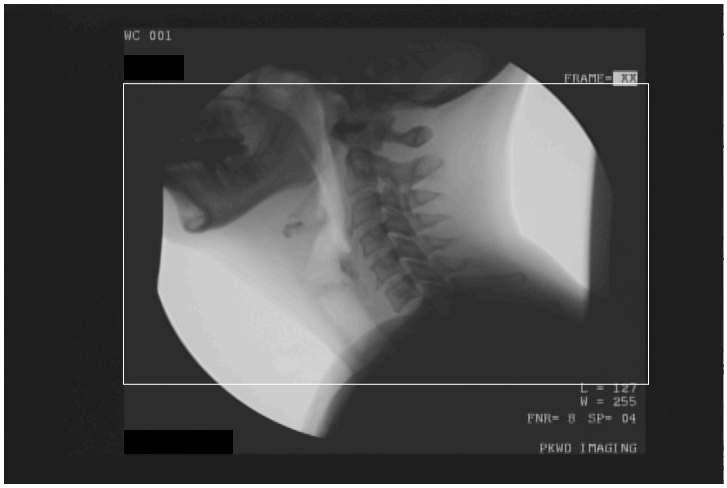


Figure 3.7: Example of the *main-roi* in a frame

3.3.2. Locating first frame for processing

As mentioned in Section 3.2.2, the swallowing process starts with a group of *number-frames* followed by a group of *XX-frames*. The processing steps associated with a particular swallow start at the first *XX-frame*. However, the first 25 *XX-frames* of the swallowing (i.e., the first second) are ignored because of the transition of the

illumination intensity of the light during this time. Figure 3.8 shows an example of a transitional intensity frame in the first second.

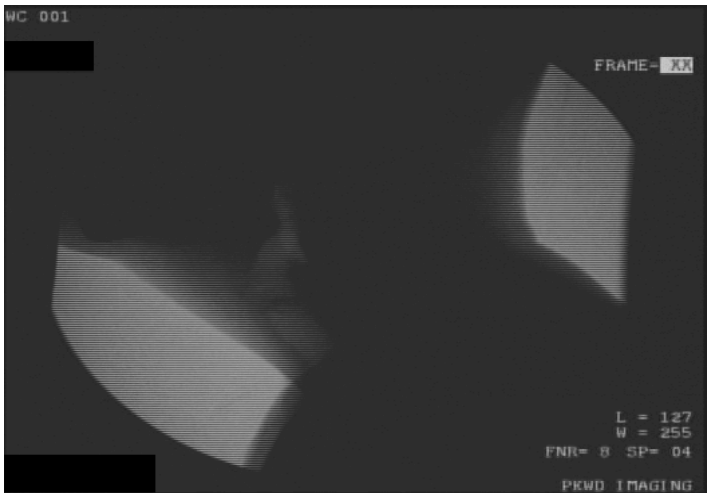


Figure 3.8: Example of a transitional intensity frame in the first second of the swallowing

3.3.3. Identifying the *Region Of Interest* (ROI)

The first step for processing data is identifying the region of interest (ROI) that contains the second, third, and fourth cervical vertebrae (C2, C3, and C4, respectively). We can refer to this area as *vertebrae-area.* Identifying this area is important because the hyoid bone and the epiglottis area are located at the left side of the *vertebrae-area*. Using the *vertebrae-area* the three cervical vertebrae (C2, C3, and C4) can be identified. In addition, the *vertebrae-area* can be used to generate a frame reference point that is anchored by the patient's location in the frame rather than the frame itself. The vertical axis (V-axis) of this reference co-ordinate is defined by the bottom-left corners of a simple bounding box around the regions identified by the C2, C3, and C4 vertebrae (see the blue and red boxes in Figure 3.9). The horizontal axis (U-axis) is

simply defined by the line perpendicular to the V-axis that passes through the bottom-left corner of the C4 box (the origin). This axis definition is illustrated in Figure 3.9.

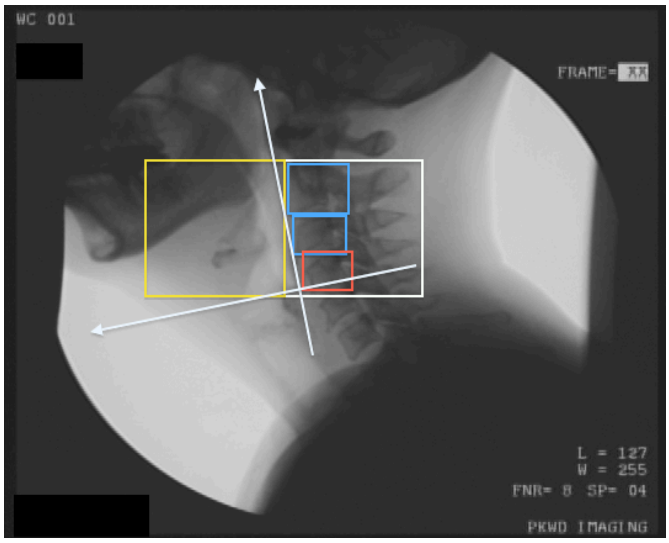


Figure 3.9: The ROI and axes defined by the patient's location in the frame. The white box is the *vertebrae-area* that contains C2, C3, and C4 boxes. The yellow box contains the hyoid bone and epiglottis.

3.3.3.1. Identifying the *vertebrae-area*

As mention in [21], the vertebrae area cannot be easily identified using typical image segmentation techniques such as edge detection, binarization and corner detection algorithms. In this context, SURF is used to identify the *vertebrae-area*.

SURF (*Speed Up Robust Features*) is a computer vision algorithm presented by Bay *et al*. in 2008 [7]. SURF is a scheme that can find corresponding points in two images for the same object. Finding the corresponding points in separate images is a three-step process:

1. Detecting interest points at distinctive locations such as corners, blobs.

2. Representing the neighborhood of each distinctive point as a feature vector called a descriptor.

3. Matching descriptors for both images based on the distance between vectors.

The OpenCV library provides a function for extracting SURF data from an image. This function can be used for object tracking and localization [2]. See Appendix B.

For this research, the SURF function is used for each frame to identify the *vertebrae-area* in the image. To identify the *vertebrae-area,* the user must select the *vertebrae-area* in the first frame to be processed. This manually selected area is used as a template image (an example is shown in Figure 3.10). Then, SURF tracks the *vertebrae-area* in the next images, which have points of interest similar to the template image.

Figure 3.10: An example of the selected vertebrae-area

The *main-roi* is used as a searching area for locating the *vertebrae-area*. SURF is rotation-invariant; that is, when the patient tilts her or his head back and forth, the SURF can still detect the *vertebrae-area*. The output of the SURF algorithm is a top left corner of the *vertebrae-area* in a current frame. Then, a box around the *vertebrae-*

area is generated with the same size and orientation as the user specified in the first frame. The generated box should follow two rules:

- The generated box should be inside the *main-roi* area.
- The maximum horizontal displacement for the *vertebrae-area* between two consecutive frames should not exceed three pixels; likewise vertical displacement should not exceed two pixels. See Section 3.3.3.3 for a more detailed discussion of this constraint.

Figure 3.11 shows an example of a frame after applying the SURF scheme, where the *vertebrae-area* is correctly identified. One advantage of using the SURF scheme is that the object can be located without training data.

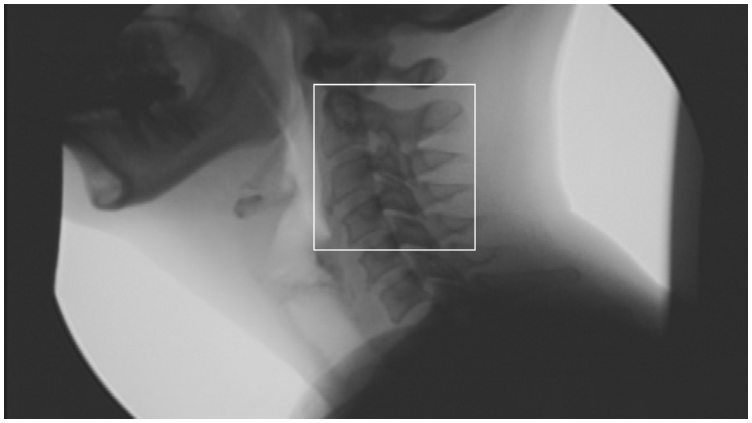


Figure 3.11: Example of a frame after applying SURF

3.3.3.2. Identify the three cervical vertebrae (C2, C3 and C4)

For tracking the three cervical vertebrae, a template-matching method is used instead of SURF. SURF does not give good results with these three cervical vertebrae because

the template images for the three cervical vertebrae are small and do not have enough points of interest to find them in *vertebrae-area*.

The OpenCV library provides a function called cvMatchTemplate that applies an appropriate template-matching technique. This function slides through the image and compares the overlapped patches with the template image, returning the patch that best matches the template image [2]. See Appendix B.

The user must individually select the three cervical vertebrae (C2, C3, and C4) in the first frame to be processed. Each of the three selections is used as a template image (an example is shown in Figure 3.12). Then, the template-matching algorithm is applied three times (once for each vertebra). The *vertebrae-area* is used as a searching area to locate the three vertebrae.

Unluckily, the template-matching scheme in OpenCV does not provide object rotation. To solve this problem, the template image is rotated 2 degrees in both directions, starting at the rotation angle from the pervious frame, each time for all frames. Then, the best match between the rotated templates and the object in the image is considered by calculating the correlation coefficient. When applying the template-matching scheme, the output is a top left corner of the cervical vertebra area. A box around the cervical vertebra is generated with the same size and orientation as the user specified in the first frame.

(a) (b) (c)

Figure 3.12: Example of the three selected vertebrae areas; (a),(b), and(c) showing the C2, C3 and C4 areas, respectively

The ideal arrangement locations for the three cervical vertebrae is shown in Figure 3.13.

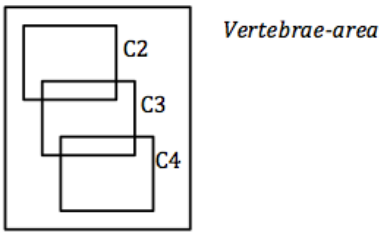


Figure 3.13: Ideal arrangement for the three cervical vertebrae

The locations of the three cervical vertebrae should satisfy several rules:

- Each box of the cervical vertebra should be inside the *vertebrae-area*. See Figure 3.14 (a).
- C2 box should be above C3 box and C3 box should be above C4 box. See Figure 3.14 (b).
- The horizontal displacement between the three cervical vertebrae boxes should be consistent. See Figure 3.14 (c).

- The vertical space between the three vertebrae should be fixed in all frames because the bone structure does not change during the swallowing process. Meaning, the space between C2-C3 or between C3-C4 should be similar in all frames.

- The maximum horizontal displacement of a vertebra between two consecutive frames should not exceed three pixels; likewise vertical displacement should not exceed two pixels. See Section 3.3.3.3 for a detailed discussion of this constraint.

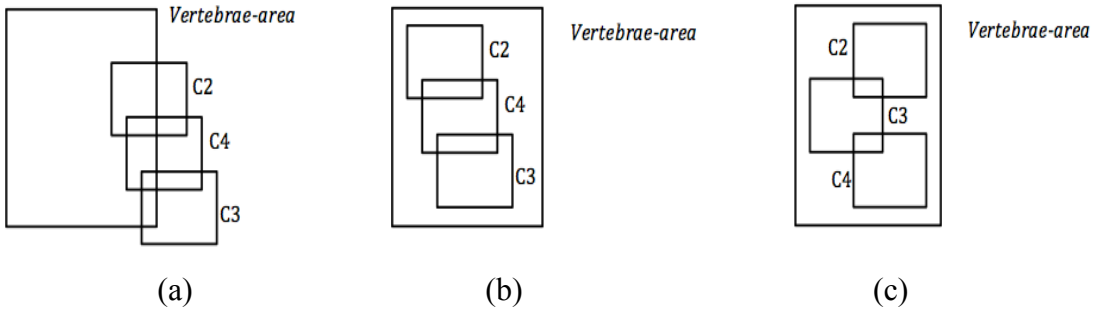


Figure 3.14: Some examples for the rejected vertebrae cases (a) vertebrae outside *vertebrae-area* (b) vertebrae in wrong order (c) vertebrae have inconsistent horizontal displacement

3.3.3.3. Determining maximum speed of patient movement

Estimating the upper bound on the speed of a patient's head and neck movement was needed to calculate several of the constraints described in Section 3.3.3.1 and Section 3.3.3.2. To do so, an experiment was performed by choosing five single-swallow videos from five different patients that exhibited noticeable movement of the patient's head. The bottom left corner of the C2 vertebra was located manually in all frames.

Then, the difference between the two locations in each pair of consecutive frames was calculated. Figure 3.15 and Figure 3.16 show the horizontal and vertical displacements for the five patients, respectively. The maximum horizontal displacement was three pixels from frame to frame (except in very few outlier cases, where the maximum was 5 pixels), while the maximum vertical displacement was only two pixels. In this research, we decided to ignore the outlier cases. Hence, we estimated the maximum horizontal displacement to be three pixels, while the maximum vertical displacement to be two pixels.

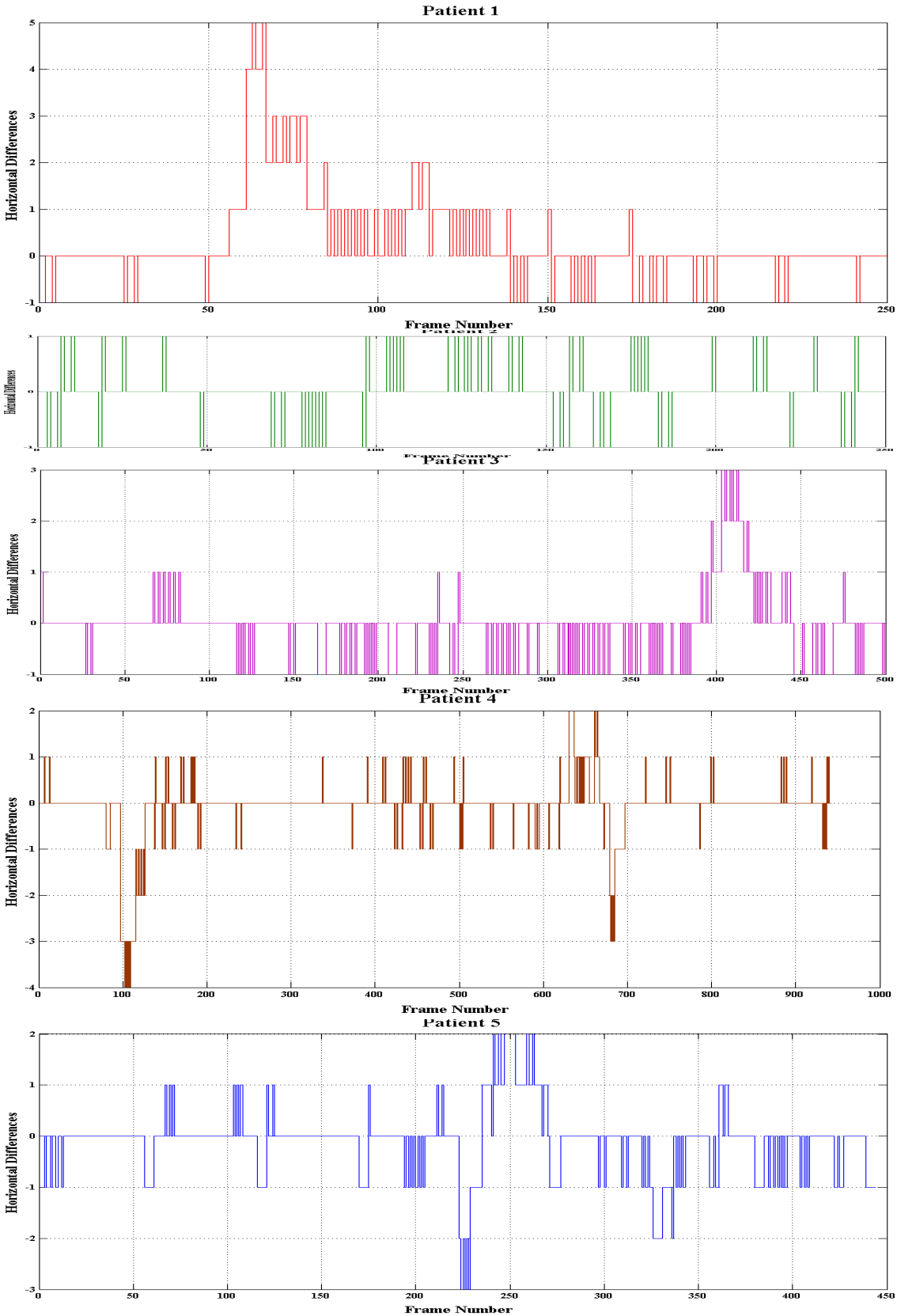


Figure 3.15: Horizontal displacement in pixels for five patients

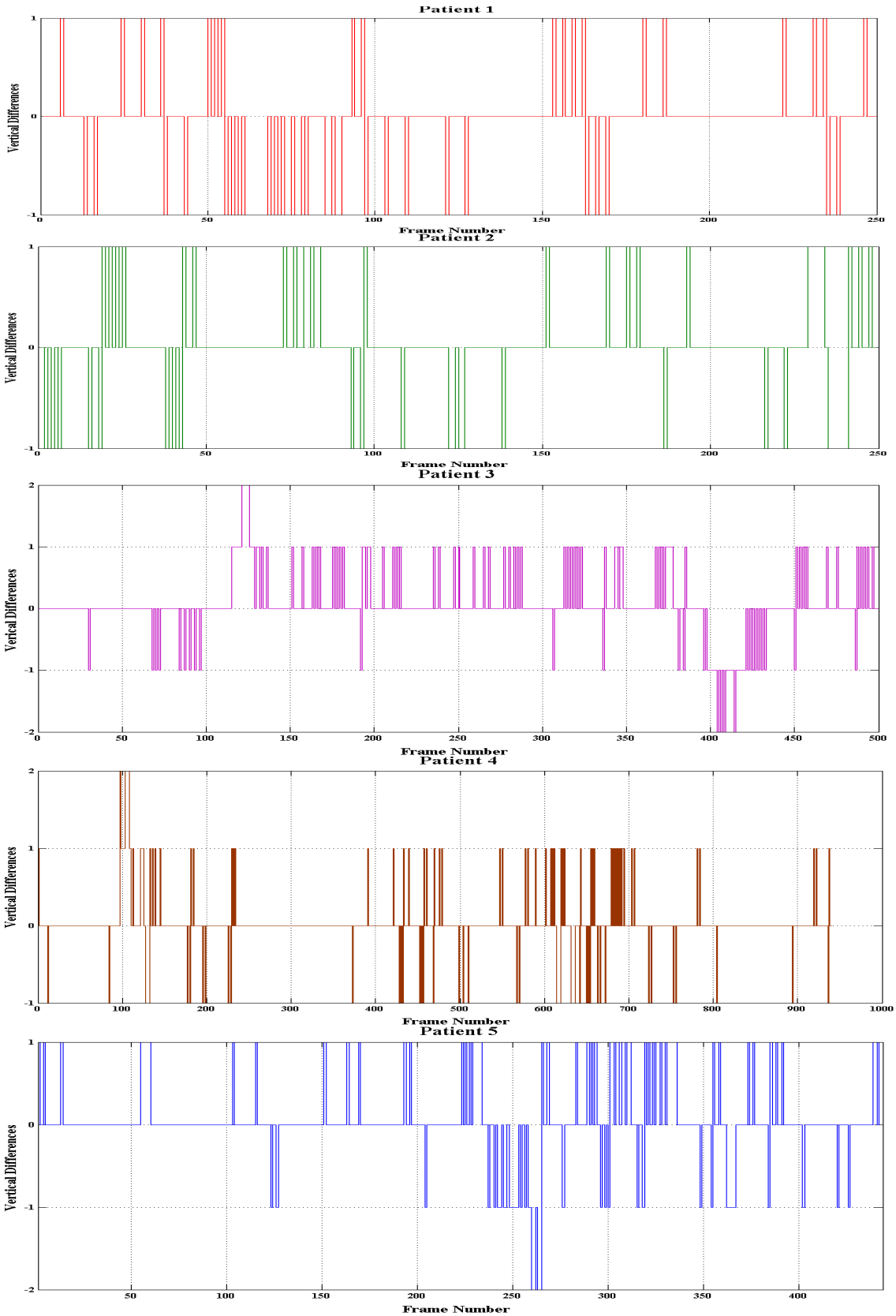


Figure 3.16: Vertical displacement in pixels for five patients

3.3.4. Generating patient-relative frame of reference

After applying the template-matching and identifying the three cervical vertebrae, the three bottom left corners of the cervical vertebrae boxes are used to generate the new axes. These three points are sent to the cvFitLine function [2] (See Appendix B) to calculate the best fitting line between them, which is the vertical axis (V) of the new axes. Then, the perpendicular line (U) that passes through the left bottom corner of the C4 can be computed. Figure 3.17 shows the simulation for the new axes and their signs. Because the new axes are relative to the cervical vertebrae of the patient, they move with the patient's movement and the final trajectory of the hyoid bone and epiglottis can be calculated relative to the patient. The patient-relative frame of reference is essential to isolate the movement of the hyoid bone and epiglottis from the patient's movement.

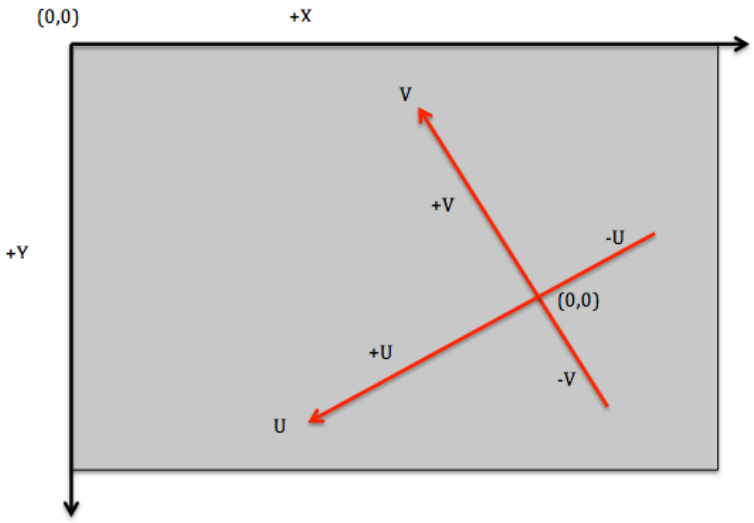


Figure 3.17: A simulation example of the new axes and their signs. The top left quarter is the positive quarter

3.3.5. Hyoid bone and epiglottis area

The area that has the hyoid bone and epiglottis is in the left side of the *vertebrae-area* (see, for example, Figure 3.9). This area is easy to identify because the *vertebrae-area* has already been identified using methods described in Section 3.3.3.1.

3.3.5.1. Tracking the hyoid bone

To accurately track the hyoid bone and the epiglottis, the user must identify them in the first frame using three points. These three points are (from left-to-right) the center of the hyoid bone, the base of the epiglottis and the tip of the epiglottis (see Figure 3.18). The optical flow Lucas-Kanade algorithm [30] is used to track the hyoid bone point across all frames.

Figure 3.18: An example of identified hyoid bone, base and tip of the epiglottis in a videofluoroscopic frame

Optical flow is a method for estimating the motion of an object in sequential frames without any prior knowledge other than its initial location. One of the most popular optical flow tracking techniques is the Lucas-Kanade (LK) algorithm. The Lucas-

Kanade algorithm was proposed in 1981 [30]. The LK algorithm relies on local information that is derived from a small window surrounding the points of interest.

The OpenCV library provides a function that calculates the optical flow using the iterative Lucas-Kanade method. This function calculates the coordinates of the points of interest on the current video frame given their coordinates on the previous frame [2]. See Appendix B.

The optical flow function receives as input the hyoid area from the previous and current frame, as well as the coordinates of the hyoid bone in the previous frame. The optical flow function returns the coordinate of the new location for the hyoid bone in the current frame. The new coordinate of the hyoid bone location should be inside the provided hyoid area.

3.3.5.2. Tracking the tip of the epiglottis

The tip of the epiglottis goes down while the hyoid bone moves upward and forward. Using the hyoid bone location, the area that has the tip of the epiglottis can be identified in order to find the exact location for the tip.

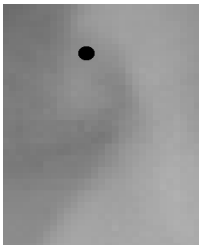


Figure 3.19: Epiglottis tip area that can be identified using hyoid bone location

To track the tip of the epiglottis, optical flow LK is applied as mentioned in Section 3.3.5.1. The optical flow function receives as input the tip area from the previous and current frame, as well as the coordinates of the tip in the previous frame. The optical flow function returns the coordinate of the new location for the tip in the current frame. To accurately locate the tip location, the tip area needs to be filtered by a Laplacian filter before applying the optical flow LK algorithm. The new tip location should be located inside the tip area and the horizontal displacement for the tip location between two consecutive frames should not exceed three pixels; likewise the vertical displacement should not exceed two pixels, as described in Section 3.3.3.3.

3.3.5.3. Tracking the base of the epiglottis

The hyoid bone location, which was identified by the LK optical flow scheme, and the tip area are used to approximate the new location for the base of the epiglottis.

As mentioned in Section 2.4.3, the first movement of the epiglottis is a consequential movement. This movement happens synchronously with the hyoid bone movement, i.e. the hyoid bone vertical and horizontal movements lead to the epiglottis movement to the transverse position. When the hyoid bone moves, the base of the epiglottis is pulled because the base of the epiglottis is attached to the ligaments between the hyoid bone and epiglottis. At the same time, the space between the base and the tip of the epiglottis are fixed because the epiglottis size does not change while swallowing.

3.3.5.3.1. Manual experiments

To demonstrate the pervious observation, manual experiments were done on 10 single-swallow videos for 10 different patients. The hyoid bone, the base and the tip of the epiglottis are located manually in all frames. The relations between the coordinates of the three points were analyzed. The correlation between horizontal and vertical displacements between hyoid bone, base and tip of the epiglottis were calculated as shown in Table 3.3.

Table 3.3: Pairwise horizontal and vertical correlation between hyoid bone, base and tip of the epiglottis

	Horizontal (X values)			Vertical (Y values)		
Patient	(H,B)	(H,T)	(B,T)	(H,B)	(H,T)	(B,T)[4]
1	0.927	-0.194	-0.032	0.945	0.061	0.194
2	0.931	0.131	0.348	0.955	0.398	0.419
3	0.915	0.094	0.398	0.975	0.709	0.615
4	0.801	0.282	0.736	0.947	0.156	0.167
5	0.924	0.548	0.688	0.862	0.593	0.574
6	0.886	0.443	0.591	0.924	0.622	0.419
7	0.952	0.857	0.94	0.899	0.682	0.565
8	0.764	0.085	0.277	0.905	0.646	0.451
9	0.925	-0.156	0.003	0.833	0.005	-0.399
10	0.777	0.257	0.266	0.961	0.899	0.862
Average	**0.880**	0.234	0.421	**0.920**	0.477	0.386

As seen in Table 3.3, the correlation coefficients between the hyoid bone and base of the epiglottis are high enough to depict a strong relation between them. When the hyoid bone moves vertically and horizontally, the base of the epiglottis moves with a

[4] H: Hyoid bone, B: Base of the epiglottis, T: Tip of the epiglottis

similar displacement. To further investigate this result, each video, is used in the previous experiment, is divided into two groups of frames. The first group has frames of the actual swallowing process; while the second group contains frames without the swallowing process (i.e. during chewing, breathing or after swallowing). The correlation coefficient for the hyoid bone and the base of the epiglottis horizontally and vertically are calculated for both groups separately, as given in Table 3.4.

Table 3.4: Pairwise correlation coefficient between the hyoid bone and the base of the epiglottis for swallowing and non-swallowing frames

	Swallowing		Non-swallowing	
Patient	Horizontal (x)	Vertical (y)	Horizontal (x)	Vertical (y)
1	0.951	0.969	0.812	0.619
2	0.902	0.936	0.938	0.91
3	0.966	0.973	0.987	0.768
4	0.788	0.907	0.895	0.767
5	0.937	0.592	0.889	0.915
6	0.864	0.923	0.982	0.923
7	0.874	0.931	0.962	0.954
8	0.787	0.902	0.738	0.911
9	0.935	0.591	0.885	0.931
10	0.715	0.371	0.55	0.974
Average	**0.871**	**0.809**	**0.863**	**0.867**

Figure 3.20 and Figure 3.21 show the horizontal and vertical displacements for the hyoid bone, base and tip of the epiglottis for the first patient. The figures show, the sold blue line (hyoid bone) and the dotted green line (base of the epiglottis) have nearly synchronized movements. The other figures that show the horizontal and vertical displacements for the rest of patients can be found in the Appendix A.

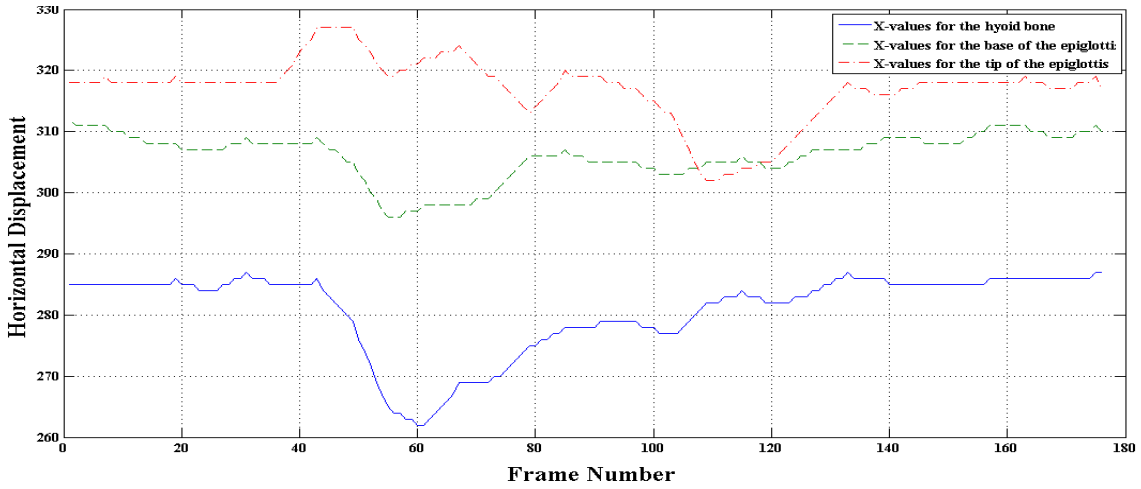


Figure 3.20: Horizontal displacement for the hyoid bone, base and tip of the epiglottis for patient 1

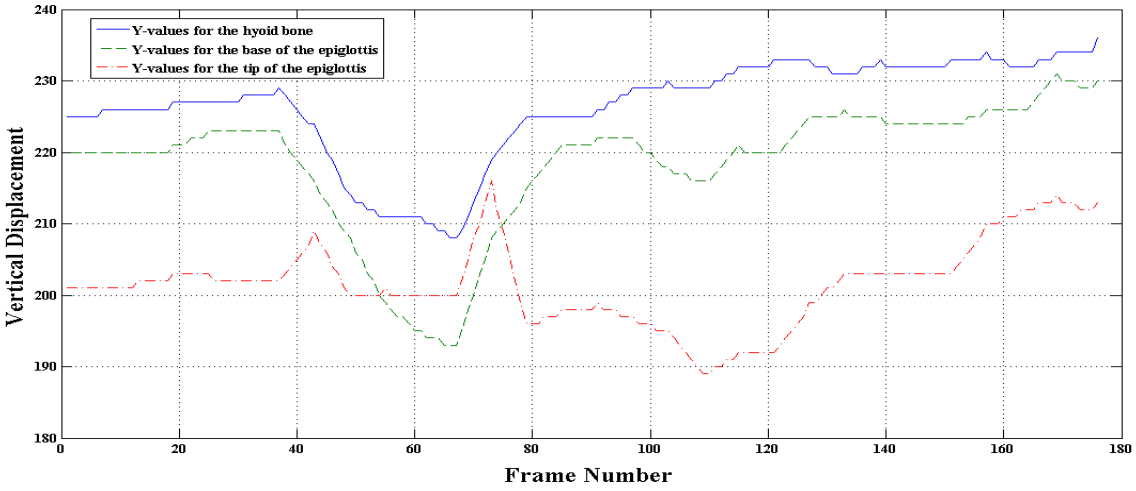


Figure 3.21: Vertical displacement for the hyoid bone, base and tip of the epiglottis for patient 1

3.3.5.3.2. Correlate the movements to find base location

The previous experiments and figures show that the hyoid bone trajectory has a correlated movement with the base of the epiglottis trajectory. Therefore, when the hyoid bone moves vertically or horizontally, the base of the epiglottis moves in the same direction. Moreover, the space between the hyoid bone and epiglottis is a fixed

value for each patient. Using this knowledge, the location of the base of the epiglottis can be calculated depending on the location of the hyoid bone and the tip of the epiglottis (see for example Figure 3.22). The base of the epiglottis should be in the tip area and should be a fixed distance from hyoid bone. In addition, the horizontal and vertical displacements of the base of the epiglottis should imitate the horizontal and vertical displacement for the hyoid bone.

Note that the location of the base of the epiglottis is computed depending on the locations of the hyoid bone and tip of the epiglottis because the base area does not have any features to track. In addition, the base of the epiglottis moves by the ligaments that are located between the hyoid bone and epiglottis. The base of the epiglottis is essential in calculating the angle of the epiglottis rotation.

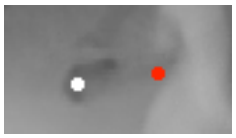


Figure 3.22: Example of the results for calculating the base of the epiglottis using the hyoid bone location. Red point represents the base of the epiglottis while the white point represents the hyoid bone location.

3.3.5.3.3. Transformed coordinates

All points that might be used for tracking the hyoid bone and epiglottis must be translated into the coordinate system of the patient-relative frame of reference. After the hyoid bone, the base, and the tip of the epiglottis locations have been determined

their coordinates are transformed to the new relative-patient reference in each frame as described in Section 3.3.4.

3.3.5.4. Calculating the hyoid bone displacement

As described in Section 3.3.5.1, the trajectory of the hyoid bone can be identified in a single-swallow video. The maximum displacements of the hyoid bone horizontally and vertically, as well as the speed of the movement can be computed. In addition, this displacement and speed can be converted to unit of millimeters (mm) for the displacement and mm/second for the speed. To convert from pixels to millimeters, a Canadian penny is attached to the patient's ear is used as a reference (see Figure 3.23). The diameter of the Canadian penny in pixels was calculated manually from different patients who are having a penny in their ear. The penny diameter is 19.05 mm and it is almost equal to 44 pixels in the patient frame. For conversion from pixel to mm, the number of pixels is multiplied by 0.431.

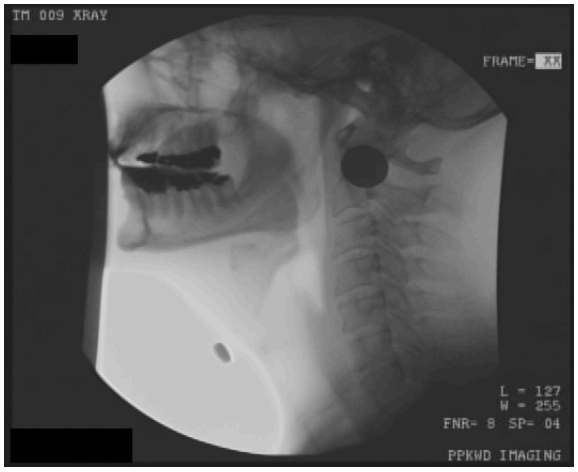


Figure 3.23: Patient that has a Canadian penny attached to his/her ear

3.3.5.5. Calculating the angle of epiglottis rotation

To calculate the angle of rotation of the epiglottis, the base and the tip of the epiglottis are tracked in all frames. Then the line between the base and the tip of the epiglottis at the resting position is considered as references for measuring the angle (see Figure 3.24). Therefore, the rotation angle can be calculated mathematically at each frame using the base and the tip locations in the current frame and the base and the tip locations in the resting position.

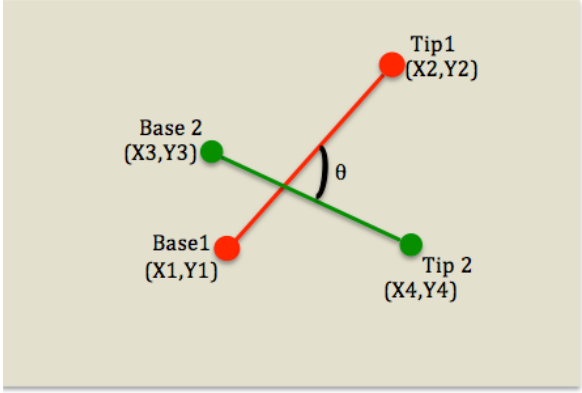


Figure 3.24: Simulation of the epiglottis angle. Base1, Tip1 are the location of the base and the tip at the resting position frame. Base2, Tip2 are the location in the other frame.

3.4. Flowchart

To summarize the described method in this chapter, two flowcharts for the pre-processing steps and the actual proposed method are shown in Figure 3.25 and Figure 3.26 respectively.

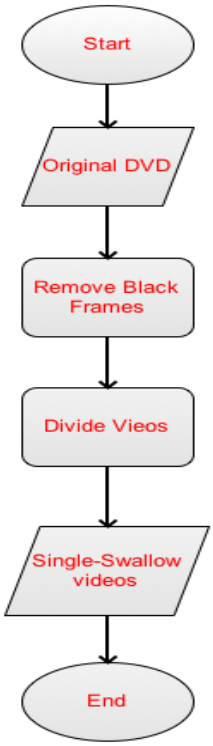


Figure 3.25: A flowchart for the pre-processing steps

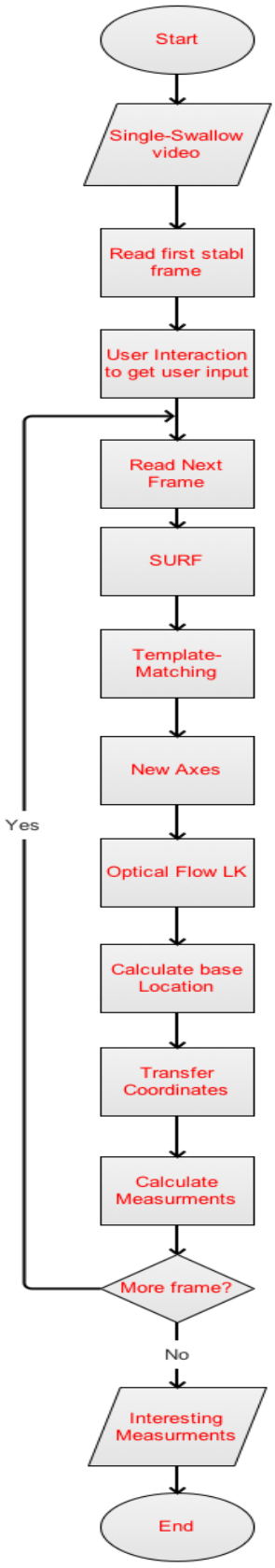


Figure 3.26: A flowchart for the proposed method

3.5. Summary

To sum up, the proposed method starts by reading a single-swallow video that is generated from the pre-processing steps. At the beginning of the proposed method, the first frame to be processed is identified automatically. Then, the region of interest that has the three cervical vertebrae (C2, C3, and C4) is identified using SURF. Each cervical vertebra is located by applying template-matching. Afterward, the hyoid bone and the tip of the epiglottis locations are tracked by optical flow. The location of the base of the epiglottis is approximated using the locations of the hyoid bone and the tip of the epiglottis. At the end, the coordinates for the hyoid bone, base and tip of the epiglottis are transformed to the new relative-patient reference. Finally, the transformed locations are used to calculate the measurements of the epiglottis and hyoid bone.

Chapter 4

Experiments and Results

This chapter describes the experiments for the proposed method that was described in Chapter 3. These experiments have been applied on videos that are recorded for patients with Parkinson's disease. Each patient swallows several kinds of food. The analysis and results are presented in the next sections.

4.1. Experiments

Each DVD goes through the three pre-processing steps that are described in Section 3.2. Then, the actual processing, which is described in Section 3.3, is applied on each single-swallow video separately. The hyoid bone, the base, and the tip of the epiglottis are tracked and their transformed coordinates are stored. Finally, the maximum displacements for the hyoid bone vertically and horizontally are calculated from the trajectory. In addition, the time and the speed that are needed to reach to the maximum horizontal and vertical places are also computed. As well, for the epiglottis, the

rotation angle in each frame as well as the maximum angle are computed using the location of the base and tip of the epiglottis.

4.2. Results for different cases

In this section, the results of three swallowing processes are displayed only. These three swallowing segments are taken from three different patients. For more results, see Appendix C.

To show the epiglottis rotation, Figure 4.1(a), Figure 4.4(a), and Figure 4.7(a) demonstrate the rotation angles in degrees in each frame for the three swallowing cases, respectively.

Figure 4.1(b), Figure 4.4(b), and Figure 4.7(b) show the horizontal displacement for the hyoid bone in millimeters for the three cases, respectively. Likewise, Figure 4.1(c), Figure 4.4(c), and Figure 4.7(c) show the vertical displacement for the hyoid bone in millimeters for the three cases, respectively.

The hyoid bone transformed coordinates are plotted in Figure 4.2(a), Figure 4.5(a), and Figure 4.8(a) after converting the units from pixels to millimeters. These plots show the trajectories of the hyoid bone in horizontal and vertical displacements for the three cases, respectively. In addition, these plots expressed only the frames that have actual swallowing process.

The final printouts from the system for the interesting measurements as described in Section 3.3.5.4 and Section 3.3.5.5 are shown in Figure 4.2(b), Figure 4.5(b), and Figure 4.8(b).

The first initial frames in each swallowing process are shown in Figure 4.3(a), Figure 4.6(a), and Figure 4.9(a) for the three cases, respectively.

One of the middle frames in the swallowing process for each case is shown in Figure 4.3(b), Figure 4.6(b), and Figure 4.9(b).

Finally, the last frames after the swallowing process for the three cases are shown in Figure 4.3(c), Figure 4.6(c), and Figure 4.9(c) respectively.

- **First Swallowing Process:**

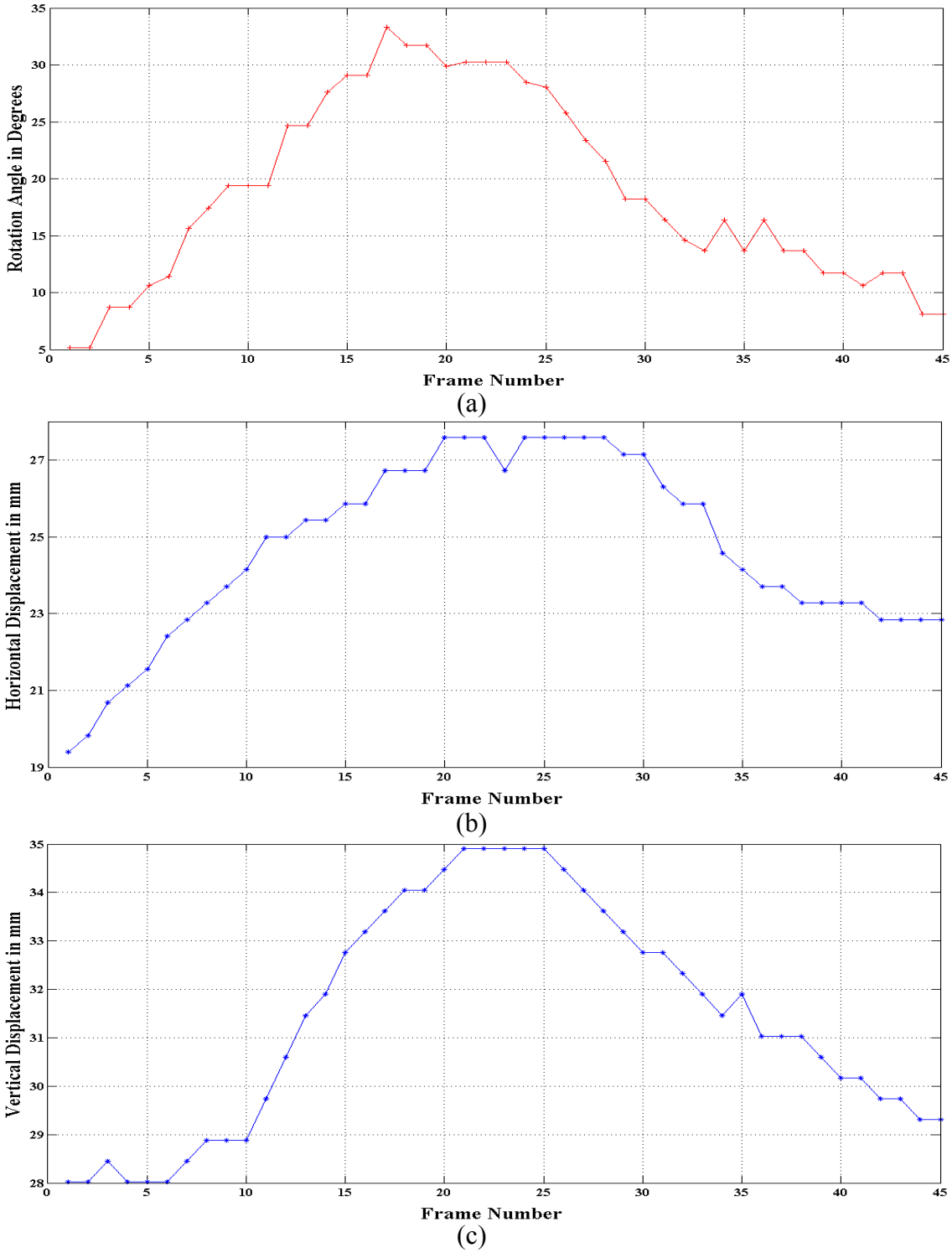
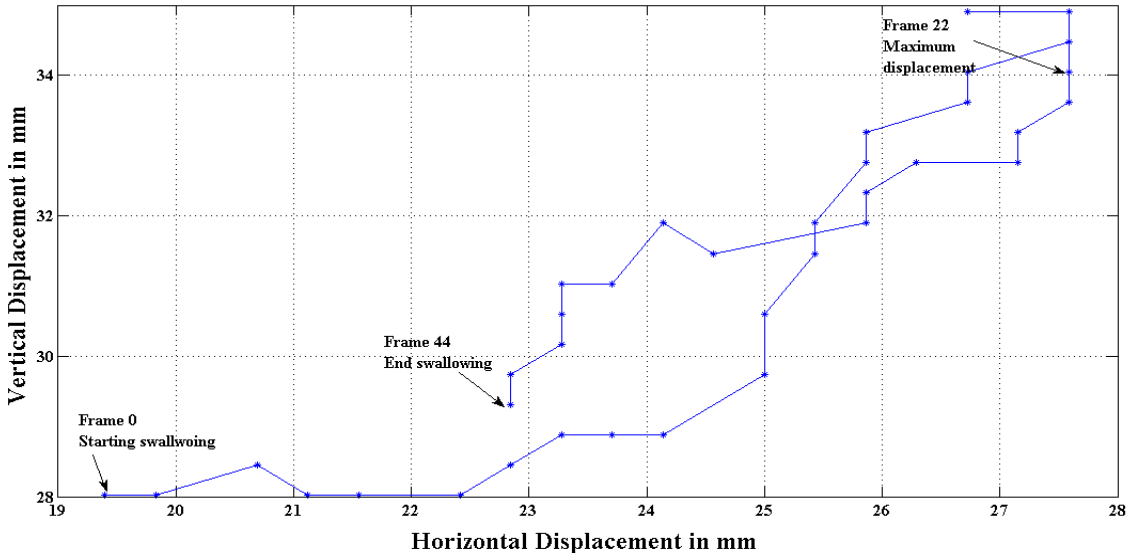


Figure 4.1: First swallowing process plots. (a) Epiglottis rotation angels in degrees. (b) Horizontal displacement of the hyoid bone in mm. (c) Vertical displacement of the hyoid bone in mm

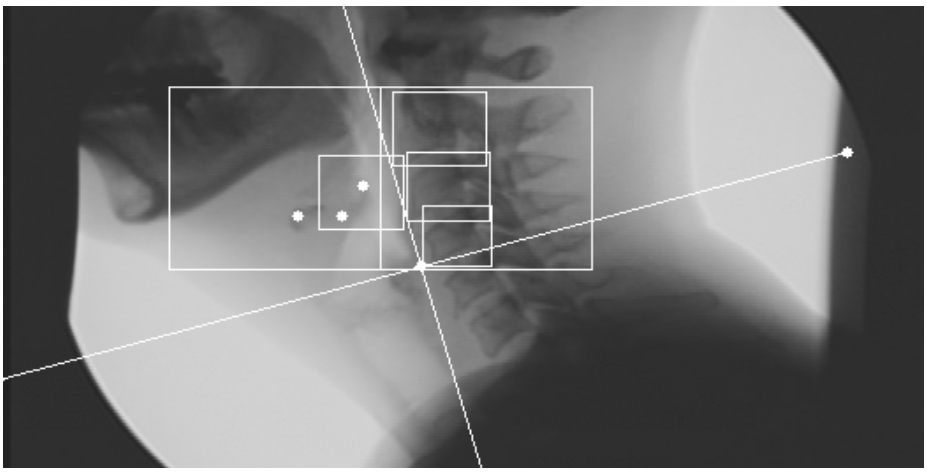


(a)

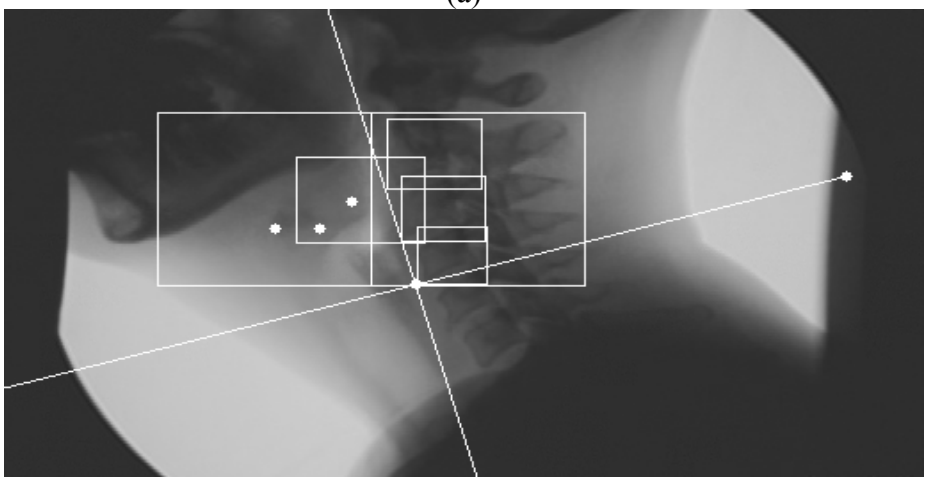
```
Hyoid bone :
 The maximum horizontal displacement = 15 pixels  equal to 6.46 mm
 The maximum vertical displacement = 20 pixels equal to 8.62 mm
 The time that need to reach the maximum horizontal displacement = 2.28 second
 The time that need to reach the maximum vertical displacement = 2.32 second
 The speed that the patient take to reach the maximum horizontal displacement = 2.84 mm/second
 The speed that the patient take to reach the maximum vertical displacement = 3.72 mm/second

 Epiglottis Rotation :
 The maximum rotation for the epiglottis = 33.33
```

(b)

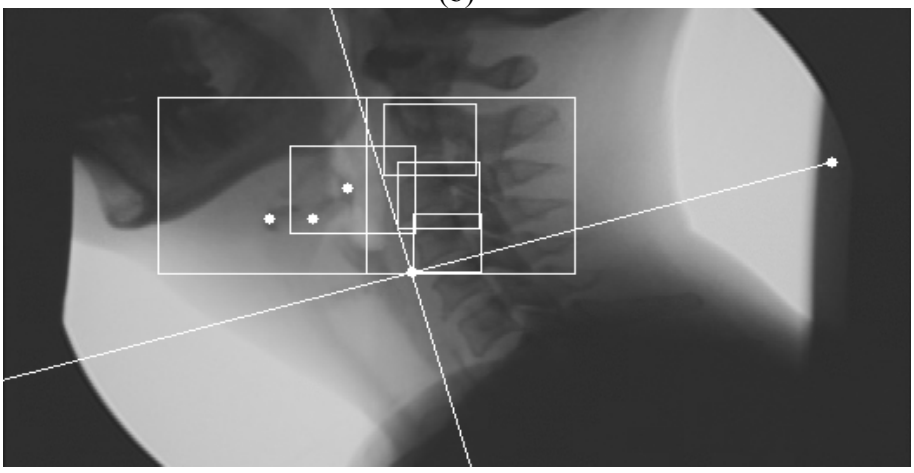
Figure 4.2: First swallowing process trajectory and printout. (a) Trajectory of the hyoid bone in mm for the first swallowing process. (b) Printout the final results from the system for this swallowing process



(a)



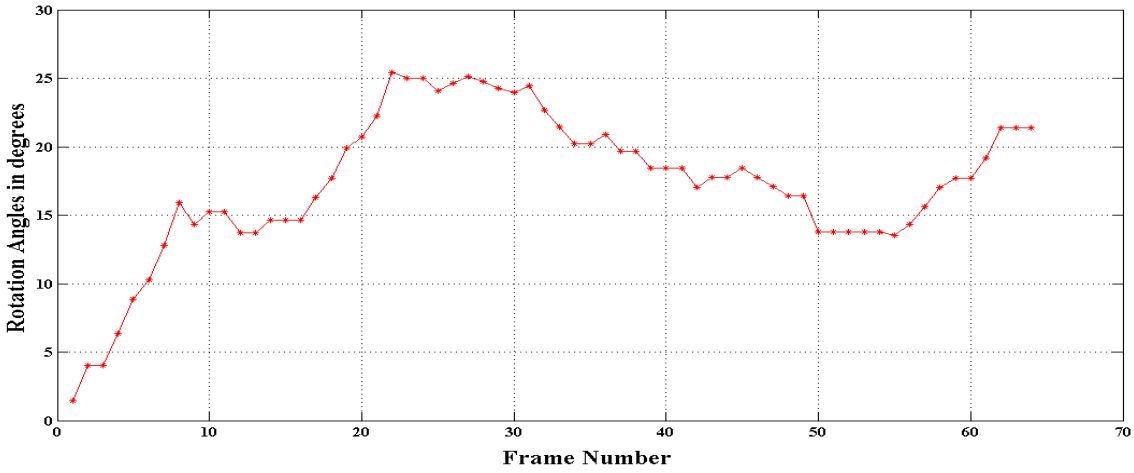
(b)



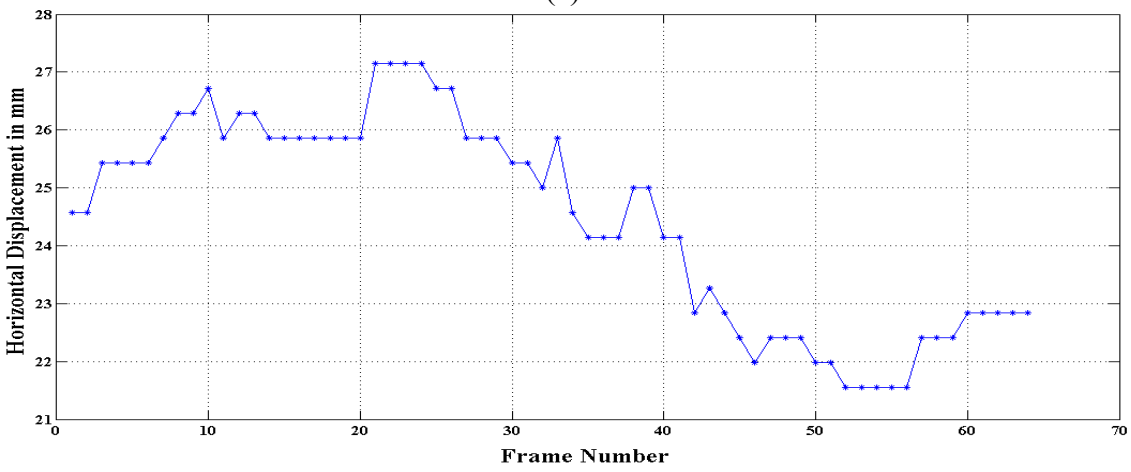
(c)

Figure 4.3: Some frames in first swallowing process. (a)First initial frame (b) Snapshot of the frame in the middle of the swallowing process. (c) Last frame in the swallowing process

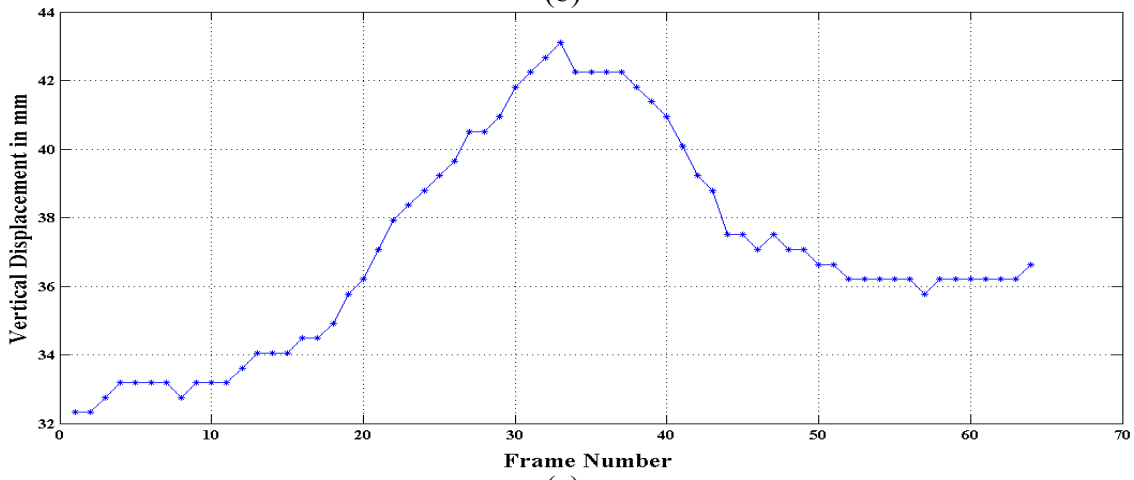
- **Second Swallowing Process:**



(a)

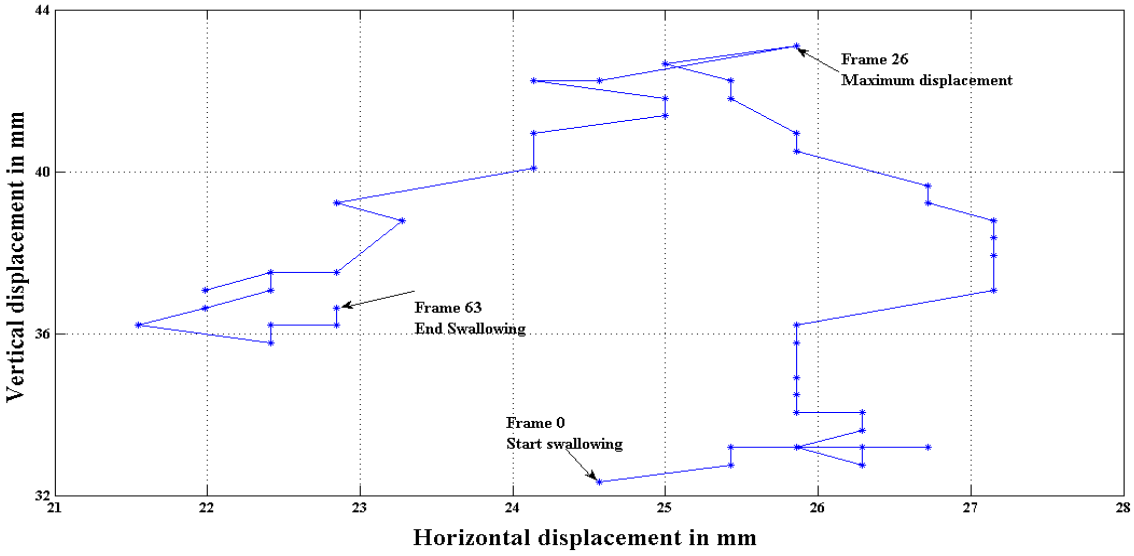


(b)



(c)

Figure 4.4: Second swallowing process plots. (a) Epiglottis rotation angels in degrees. (b) Horizontal displacement of the hyoid bone in mm. (c) Vertical displacement of the hyoid bone in mm

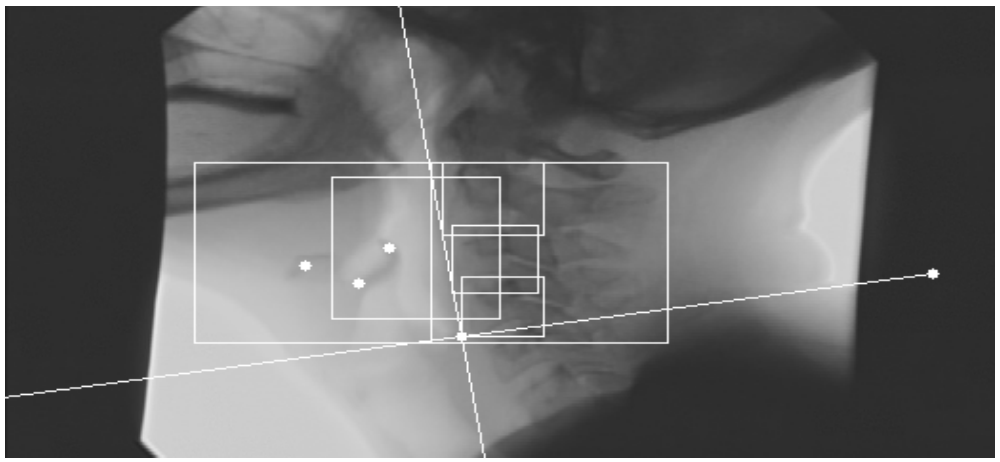


(a)

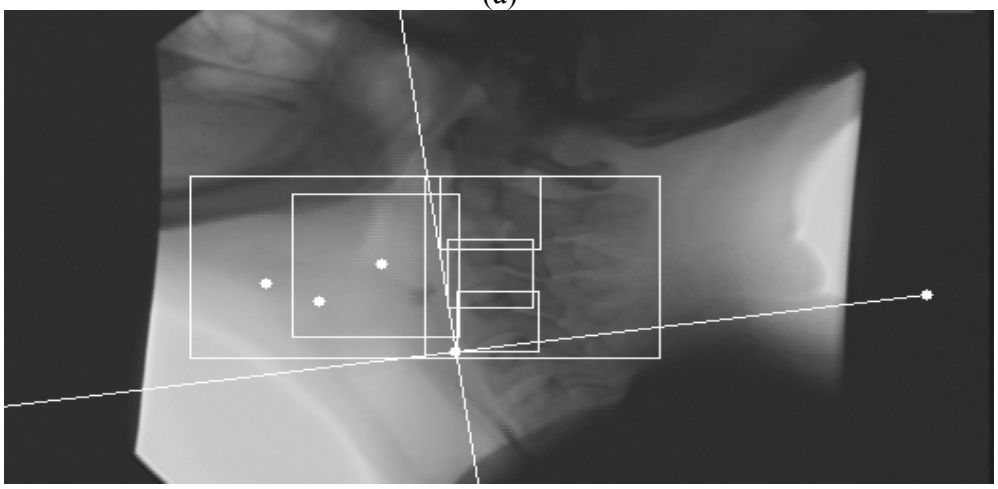
```
Hyoid bone :
The maximum horizontal displacement = 6 pixels  equal to 2.59 mm
The maximum vertical displacement = 25 pixels equal to 10.78 mm
The time that need to reach the maximum horizontal displacement = 0.80 second
The time that need to reach the maximum vertical displacement = 1.28 second
The speed that the patient take to reach the maximum horizontal displacement = 3.23 mm/second
The speed that the patient take to reach the maximum vertical displacement = 8.42 mm/second

Epiglottis Rotation :
The maximum rotation for the epiglottis = 27.12
```

(b)

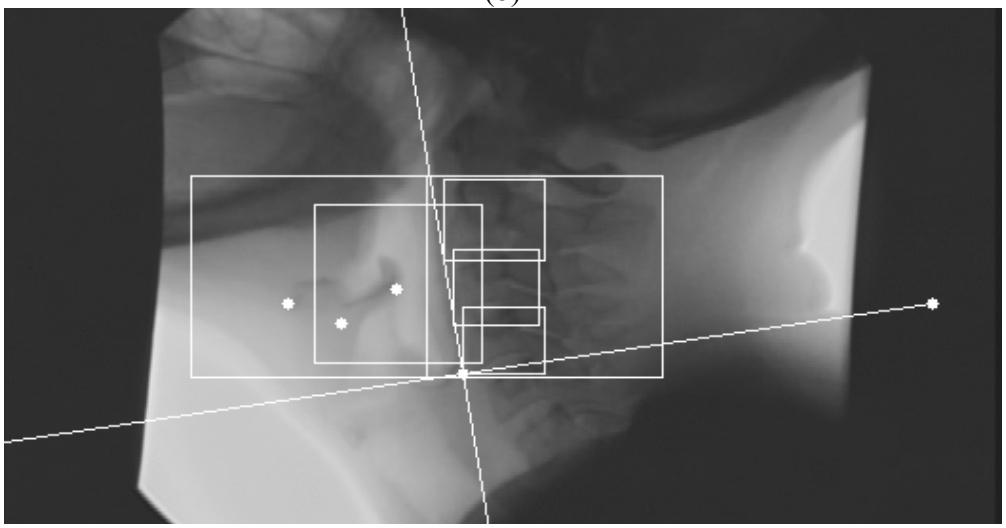
Figure 4.5: Second swallowing process trajectory and printout. (a) Trajectory of the hyoid bone in mm for the first swallowing process. (b) Printout the final results from the system for this swallowing process



(a)



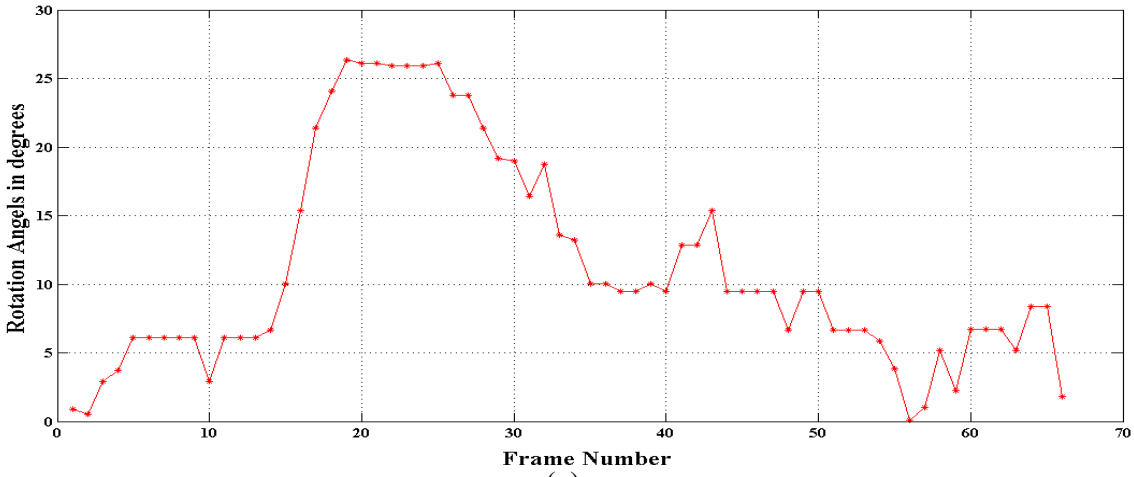
(b)



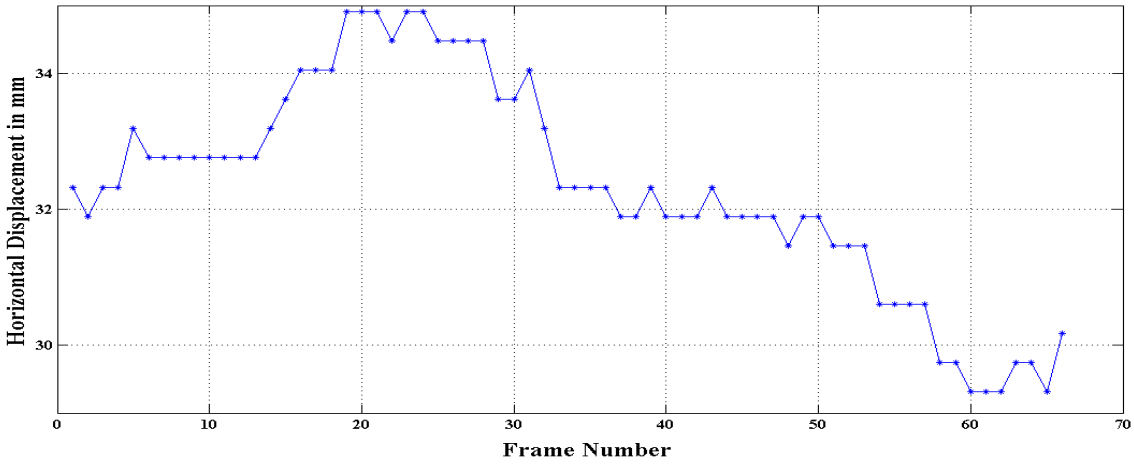
(c)

Figure 4.6: Some frames in second swallowing process. (a)First initial frame (b) Snapshot of the frame in the middle of the swallowing process. (c) Last frame in the swallowing process

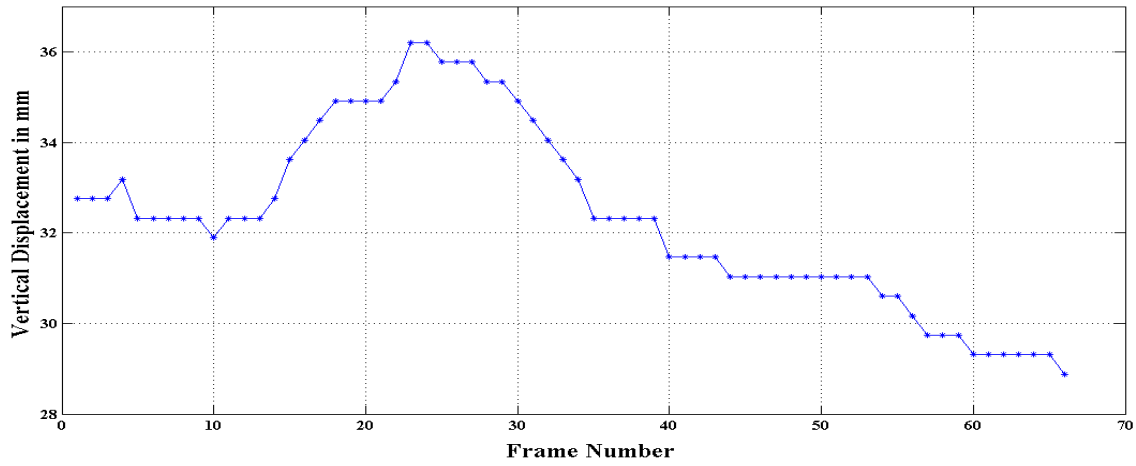
- **Third Swallowing Process:**



(a)

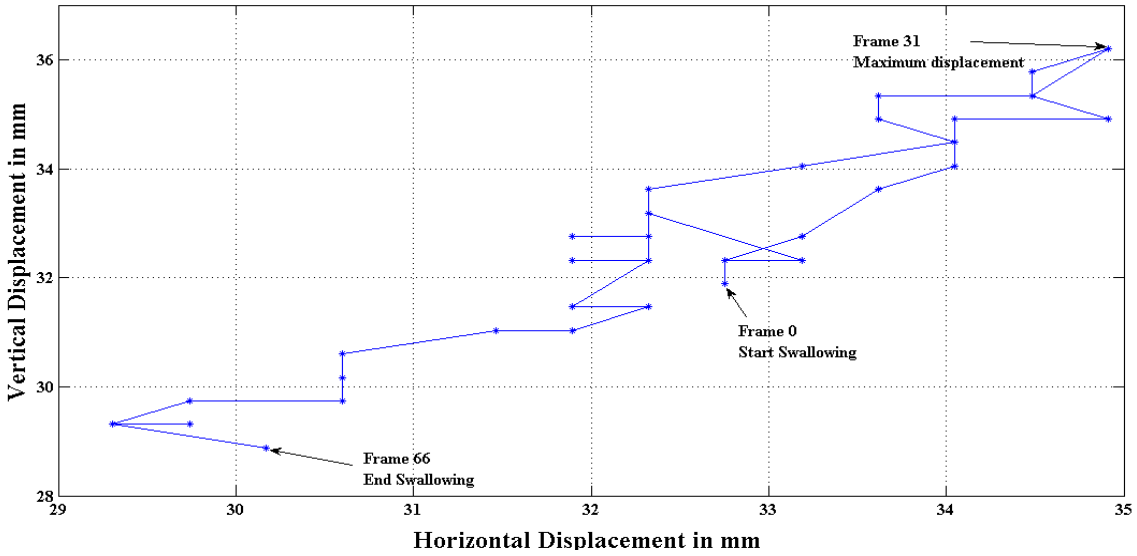


(b)



(c)

Figure 4.7: Third swallowing process plots. (a) Epiglottis rotation angels in degrees. (b) Horizontal displacement of the hyoid bone in mm. (c) Vertical displacement of the hyoid bone in mm

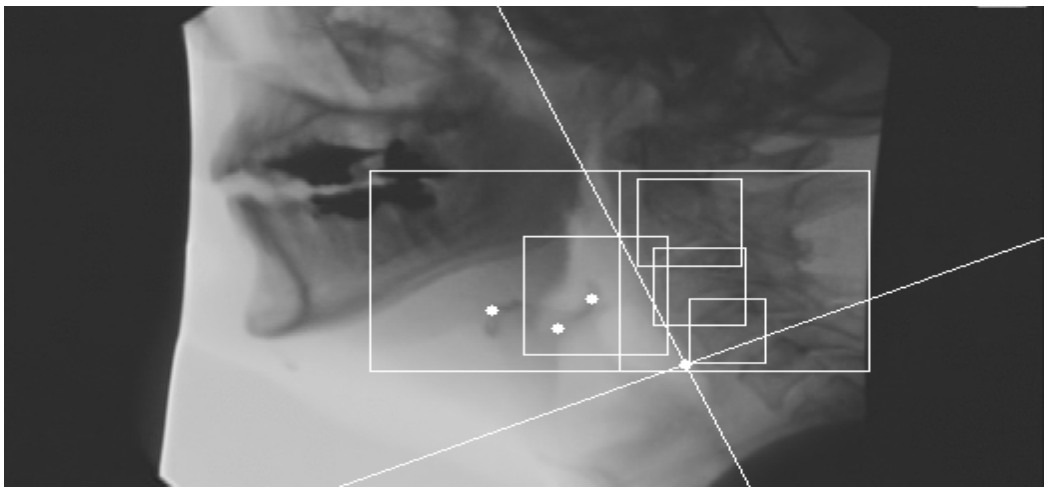


(a)

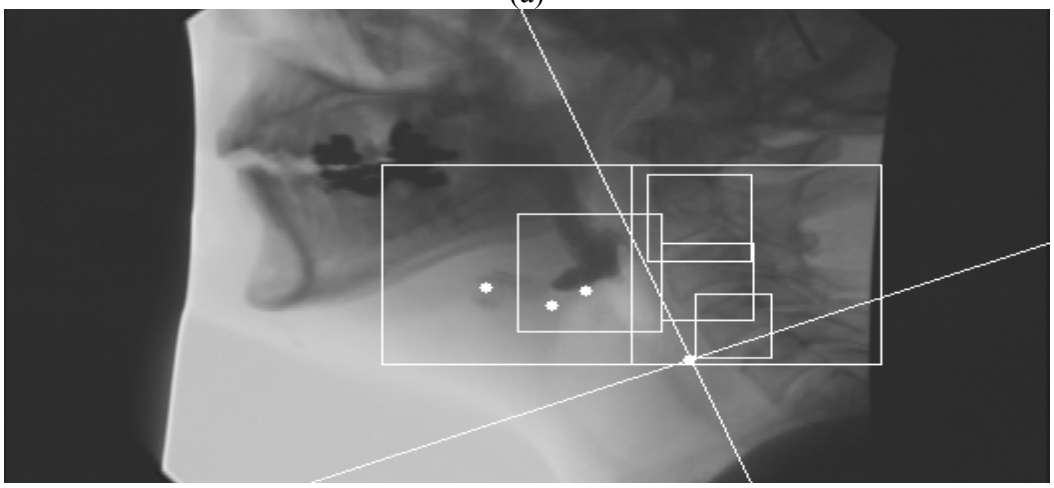
```
Hyoid bone :
The maximum horizontal displacement = 11 pixels  equal to 4.74 mm
The maximum vertical displacement = 8 pixels equal to 3.45 mm
The time that need to reach the maximum horizontal displacement = 1.16 second
The time that need to reach the maximum vertical displacement = 1.32 second
The speed that the patient take to reach the maximum horizontal displacement = 4.09 mm/second
The speed that the patient take to reach the maximum vertical displacement = 2.61 mm/second

Epiglottis Rotation :
The maximum rotation for the epiglottis = 26.34
```

(b)

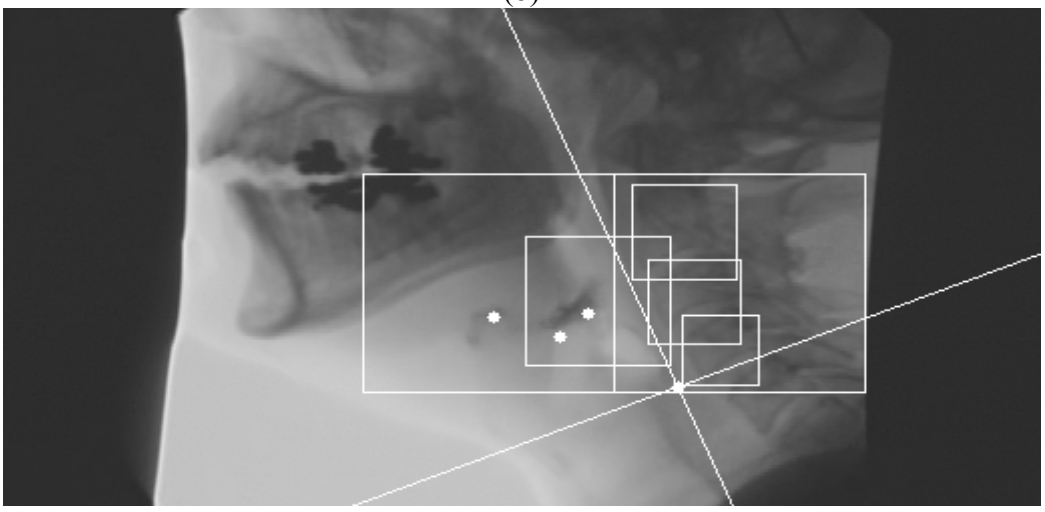
Figure 4.8: Third swallowing process trajectory and printout. (a) Trajectory of the hyoid bone in mm for the first swallowing process. (b) Printout the final results from the system for this swallowing process



(a)



(b)



(c)

Figure 4.9: Some frames in third swallowing process. (a)First initial frame (b) Snapshot of the frame in the middle of the swallowing process. (c) Last frame in the swallowing process

4.3. Analysis

As shown from the previous results in Section 4.2, the proposed CAD system can track the hyoid bone and epiglottis semi-automatically. When the hyoid bone reaches its maximum horizontal and vertical location, the epiglottis rotation angle also reaches its maximum value, i.e. when the hyoid bone moves horizontally and vertically, the epiglottis rotates down. Then, when the hyoid bone goes back, the epiglottis goes back as well to the resting position. In addition, for the hyoid bone, the trajectories show that the hyoid bone moves upward and forward at the beginning of the swallowing then it goes back to the original location, which follows the typical movement for the hyoid bone. For the epiglottis rotation, the angle increases to reach a maximum value then goes back to the original position, which is the typical movement for the epiglottis. To evaluate patient situation, a radiologists can interpolate the printout of the system that shows the measurements in actual units (mm, second, mm/second, and degree).

The proposed method and results have been verified and approved by a Speech-Language Pathologist specialist at London Health Science Center.

4.4. Summary

This chapter presents results and analysis for the proposed method in several swallowing cases. These results show the correlated relation between the hyoid bone and the epiglottis movements. In addition, one expert has verified these results.

Chapter 5

Conclusions and Future Work

5.1. Summary and conclusions

In this thesis, a CAD system that can track the hyoid bone and the epiglottis semi-automatically is proposed. This system can help the radiologists to measure the hyoid bone and epiglottis movement during swallowing. The proposed system needs small interaction from the user to indicate special area and points in the first process frame. Then, the *vertebrae-area* can be identified automatically by using SURF. Template-matching is used to identify the three cervical vertebrae (C2, C3, and C4). The hyoid bone and the tip of the epiglottis can be tracked in all frames using the LK optical flow algorithm. Then, the base of the epiglottis is computed using the hyoid bone and tip locations. At the end, the system transforms the calculated coordinates of the hyoid bone and the base and the tip of the epiglottis to the new patient-relative system.

The hyoid bone and epiglottis measurements have essential roles to evaluate the

success of swallowing. The proposed method provides results that can be used in therapeutic sessions to evaluate the patient status.

Tracking the tip of the epiglottis is a challenging process because it can be missed in some frames and its features are so weak. The lack of studies that try to track the epiglottis from x-ray images adds to the challenge. There are some researchers who studied the epiglottis movement medically. However, there is no study that tries to track it using computer vision tools.

The proposed system is considered to be a pioneer for tracking the epiglottis, as it is the first of its kind that tries to semi-automatically track the epiglottis from digital fluoroscopic images.

5.2. Limitation

Even though the proposed CAD can track the hyoid bone and the epiglottis successfully in many cases, it does not give good results for the epiglottis when the patient swallows liquid. Swallowing liquids hides the tip of the epiglottis for a long time. Therefore, the LK optical flow will track the liquid instead of the tip. In addition, the single-swallow videos might contain some blurred images. In such case, SURF and template-matching schemes cannot find the areas of interest correctly. Moreover, the proposed method assumes that the ligaments between the hyoid bone and the epiglottis are intact. If there are any problems with ligaments, the base location might be incorrect. Although some cases do not work with the system, it significantly adds to the field because, to the best of our knowledge, tracking the epiglottis semi-automatically is a new research.

5.3. Future Work

The proposed method is built to function when patients swallow a spoonful of fluid and solid food. Further research is needed to expand our system to deal with larger amounts of fluid. Also, the proposed method can be improved by including techniques to fix the blurred frames to avoid incorrect detection of areas of interest. Moreover, the method can be further improved by introducing automatic detection of the first and last swallowing frames in order to track the hyoid bone and epiglottis locations on the swallowing-only frames. In addition, for more verification, the system results might be compared with manual assessment by experts. Finally, our proposed method may be integrated with a larger system that calculates further measurements semi-automatically, which in turn, will aid radiologists in their work.

Bibliography

[1] National Institute of Neurological Disorders and Stroke. < http://www.ninds.nih.gov/disorders/parkinsons_disease/parkinsons_disease.htm#What_is. >, Last seen May, 2013

[2] OpenCV Documentation <http://docs.opencv.org/. >, Last seen May, 2013

[3] M. Aung, J. Goulermas, S. Hamdy and M. Power, "Spatiotemporal visualizations for the measurement of oropharyngeal transit time from videofluoroscopy," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 2, pp. 432-441 2010.

[4] M. Aung, J. Goulermas, S. Stanschus, S. Hamdy and M. Power, "Automated anatomical demarcation using an active shape model for videofluoroscopic analysis in swallowing," *Medical Engineering and Physics*, vol. 32, no. 10, pp. 1170-1179 2010.

[5] M. Aung, S. Hamdy, M. Power, S. Stanschus and J. Goulermas, "Measuring Bolus Transit Times from Videofluoroscopy Using Image Profiles and Particle Swarm Optimisation," in *Developments in eSystems Engineering, 2009 Second International Conference on*, 2009, pp. 117-122.

[6] L. Baijens, R. Speyer, V. Passos, W. Pilz, N. Roodenburg and P. Clave, "Swallowing in parkinson patients versus healthy controls: Reliability of measurements in videofluoroscopy," *Gastroenterology Research & Practice*, vol. 2011, pp. 1-9, 01 2011.

[7] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, "Speeded-up robust features (SURF)," *Comput.Vision Image Understanding*, vol. 110, no. 3, pp. 346-359, 6 2008.

[8] M. Ceccarelli, A. Colaprico, L. De Vito, S. Marotta and M. Piscitelli, "A semi-automatic measurement system for the swallowing analysis in videofluoroscopy," in *Medical Measurements and Applications Proceedings, 2011 IEEE International Workshop on*, 2011, pp. 125-130.

[9] Y. Chen, J. Barron, D. Taves and R. Martin, "Computer measurement of oral movement in swallowing," *Dysphagia*, vol. 16, no. 2, pp. 97-109, 03/01 2001.

[10] K. Choi, J. Ryu, M. Kim, J. Kang and S. Yoo, "Kinematic analysis of dysphagia: Significant parameters of aspiration related to bolus viscosity," *Dysphagia*, vol. 26, no. 4, pp. 392-398 2011.

[11] S. Daniels, M. Schroeder, M. McClain, D. Corey, J. Rosenbek and A. Foundas, "Dysphagia in stroke: Development of a standard method to examine swallowing recovery," *Journal of rehabilitation research and development*, vol. 43, no. 3, pp. 347-355 2006.

[12] K. Doi, "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Comput.Med.Imaging Graphics*, vol. 31, no. 4-5, pp. 198-211 2007.

[13] J. Dyer, P. Leslie and M. Drinnan, "Objective computer-based assessment of valleculae residue – is it useful?" *Dysphagia*, vol. 23, no. 1, pp. 7-15 2008.

[14] A. Eisbruch, T. Lyden, C. Bradford, L. Dawson, M. Haxer, A. Miller, T. Teknos, D. Chepeha, N. Hogikyan, J. Terrell and G. Wolf, "Objective assessment of swallowing dysfunction and aspiration after radiation concurrent with chemotherapy for head-and-neck cancer," *International Journal of Radiation Oncology, Biology, Physics*, vol. 53, no. 1, pp. 23-28, 5/1 2002.

[15] E. Eisenhuber, W. Schima, E. Schober, P. Pokieser, A. Stadler, M. Scharitzer and E. Oschatz, "Videofluoroscopic assessment of patients with dysphagia: Pharyngeal retention is a predictive factor for aspiration," *American Journal of Roentgenology*, vol. 178, no. 2, pp. 393-398 2002.

[16] O. Ekberg and S. Sigurjónsson, "Movement of the epiglottis during deglutition," *Gastrointestinal Radiology*, vol. 7, no. 1, pp. 101-107 1982.

[17] B. Garon, Z. Huang, S. Hommeyer, D. Eckmann, G. Stern and C. Ormiston, "Epiglottic dysfunction: Abnormal epiglottic movement patterns," *Dysphagia*, vol. 17, no. 1, pp. 57-68 2002.

[18] M. Gokyigit, N. Pazarci, I. Ercan, S. Seker, S. Turgut and C. Ertekin, "Identification of distinct swallowing patterns for different bolus volumes," *Clinical Neurophysiology*, vol. 120, no. 9, pp. 1750-1754, 9 2009.

[19] T. Han, N. Paik and J. Park, "Quantifying swallowing function after stroke: A functional dysphagia scale based on videofluoroscopic studies," *Arch.Phys.Med.Rehabil.*, vol. 82, no. 5, pp. 677-682 2001.

[20] R. Higo, T. Nito and N. Tayama, "Videofluoroscopic assessment of swallowing function in patients with myasthenia gravis," *Journal of the Neurological sciences.*, vol. 231, no. 1-2, pp. 45-48 2005.

[21] I. Hussain, "Semi-automatic assessment of hyoid bone motion in digital videouoroscopic images," 2012.

[22] Y. Jang, S. Lee, J. Jeon and S. Lee, "Analysis of video fluoroscopic swallowing study in patients with vocal cord paralysis," *Dysphagia*, pp. 1-6 2011.

[23] P. Kahrilas, S. Lin, J. Chen and J. Logemann, "Oropharyngeal accommodation to swallow volume," *Gastroenterology*, vol. 111, no. 2, pp. 297-306 1996.

[24] P. Kellen, D. Becker, J. Reinhardt and D. Van Daele, "Computer-assisted assessment of hyoid bone motion from videofluoroscopic swallow studies," *Dysphagia*, vol. 25, no. 4, pp. 298-306 2010.

[25] K. Kendall and R. Leonard, "Videofluoroscopic upper esophageal sphincter function in elderly dysphagic patients," *Laryngoscope*, vol. 112, no. 2, pp. 332-337 2002.

[26] K. Kluin, M. Bromberg, E. Feldman and Z. Simmons, "Dysphagia in elderly men with myasthenia gravis," *Journal of the Neurological sciences*, vol. 138, no. 1-2, pp. 49-52 1996.

[27] K. Kuhlemeier, P. Yates and J. Palmer, "Intra- and interrater variation in the evaluation of videofluorographic swallowing studies," *Dysphagia*, vol. 13, no. 3, pp. 142-147 1998.

[28] R. Leonard and P. Belafsky, "Dysphagia following cervical spine surgery with anterior instrumentation: Evidence from fluoroscopic swallow studies," *Spine*, vol. 36, no. 25, pp. 2217-2223 2011.

[29] J. Logemann, in *Evaluation and treatment of swallowing disorder*, 1983, pp. 11-26.

[30] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th international joint conference on Artificial intelligence*, 1981.

[31] G. McCullough, R. Wertz, J. Rosenbek, R. Mills, W. Webb and K. Ross, "Inter- and intrajudge reliability for videofluoroscopic swallowing evaluation measures," *Dysphagia*, vol. 16, no. 2, pp. 110-118 2001.

[32] S. Namaki, T. Tanaka, Y. Hara, H. Ohki, M. Shinohara and Y. Yonhehara, "Videofluorographic evaluation of dysphagia before and after modification of the flap and scar in patients with oral cancer," *Journal of Plastic Surgery and Hand Surgery*, vol. 45, no. 3, pp. 136-142, 06/01; 2012/02 2011.

[33] N. Paik, S. Kim, H. Lee, J. Jeon, J. Lim and T. Han, "Movement of the hyoid bone and the epiglottis during swallowing in patients with dysphagia from different etiologies," *Journal of Electromyography and Kinesiology*, vol. 18, no. 2, pp. 329-335 2008.

[34] J. Palmer, K. Kuhlemeier, D. Tippett and C. Lynch, "A protocol for the videofluorographic swallowing study," *Dysphagia*, vol. 8, no. 3, pp. 209-214 1993.

[35] A. Perlman, X. He, J. Barkmeier and E. Van Leer, "Bolus location associated with videofluoroscopic and respirodeglutometric events," *Journal of Speech, Language, and Hearing Research*, vol. 48, no. 1, pp. 21-33 2005.

[36] J. Robbins, J. Coyle, J. Rosenbek, E. Roecker and J. Wood, "Differentiation of normal and abnormal airway protection during swallowing using the Penetration–Aspiration scale," *Dysphagia*, vol. 14, no. 4, pp. 228-232 1999.

[37] L. Rofes, V. Arreola, J. Almirall, M. Cabré, L. Campins, P. García-Peris, R. Speyer and P. Clavé, "Diagnosis and management of oropharyngeal dysphagia and its nutritional and respiratory complications in the elderly," *Gastroenterology research and practice*, vol. 2011, pp. 1-13 2011.

[38] N. Rommel, E. Bellon, R. Hermans, M. Smet, A. De Meyer, L. Feenstra, E. Dejaeger and G. Veereman-Wauters, "Development of the orohypopharyngeal cavity in normal infants and young children," *The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association*, vol. 40, no. 6, pp. 606-611 2003.

[39] J. Rosenbek, J. Robbins, E. Roecker, J. Coyle and J. Wood, "A penetration-aspiration scale," *Dysphagia*, vol. 11, no. 2, pp. 93-98 1996.

[40] S. Stoeckli, T. Huisman, B. Seifert and B Martin–Harris, "Interrater reliability of videofluoroscopic swallow evaluation," *Dysphagia*, vol. 18, no. 1, pp. 53-57 2003.

[41] J. Sun, B. Xu, Y. Yuan and J. Xu, "Study on the function of pharynx & upper esophageal sphincter in globus hystericus," *World Journal Of Gastroenterology*, vol. 8, no. 5, pp. 952-955 2002.

[42] C. Tsushima, E. Saitoh, M. Baba, M. Yokoyama, W. Fujii, S. Okada and H. Uematsu, "Hyoid movement and laryngeal penetration during sequential swallowing," *Journal of Medical and Dental Sciences*, vol. 56, no. 3, pp. 113-21 2009.

[43] T. Warabi, T. Ito, M. Kato, H. Takei, N. Kobayashi and S. Chiba, "Effects of stroke-induced damage to swallow-related areas in the brain on swallowing mechanics of elderly patients," *Geriatrics & gerontology international*, vol. 8, no. 4, pp. 234-242 2008.

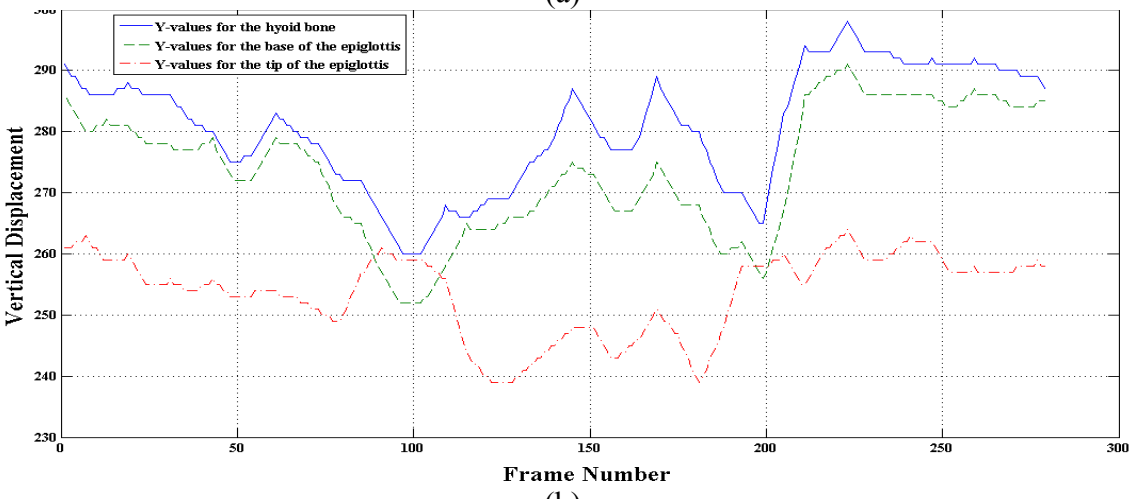
[44] K. Yabunaka, M. Ohue, T. Hashimoto, T. Katsuda, K. Yamamoto and S. Sanada, "Sonographic analysis of hyoid bone movement during swallowing," in *International Federation for Medical and Biological Engineering Proceedings*, 2008, pp. 840-842.

Appendices

Appendix A. Manual Experiments Plots

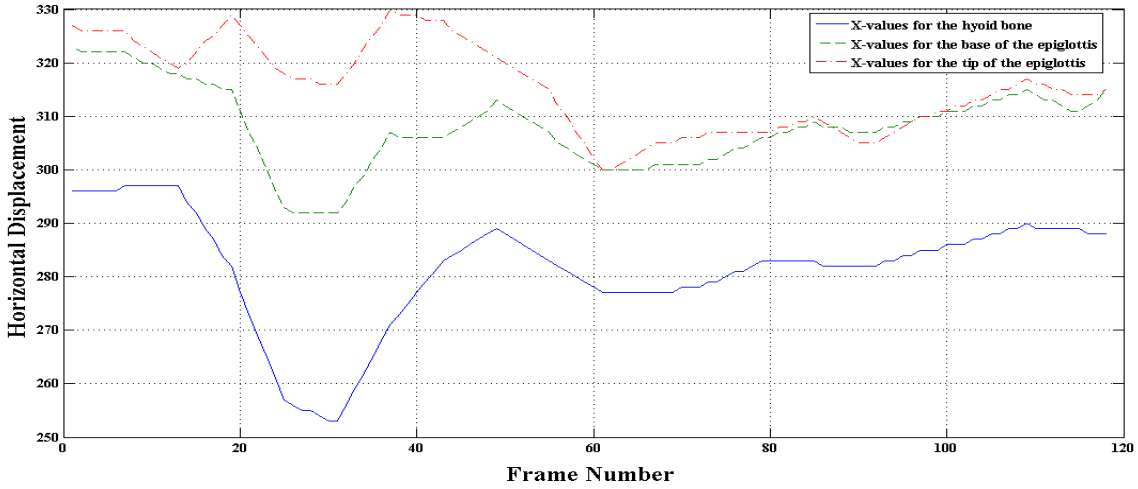
Plots for horizontal and vertical displacements for the hyoid bone, base and tip of the epiglottis that are located manually in10 patients

(a)

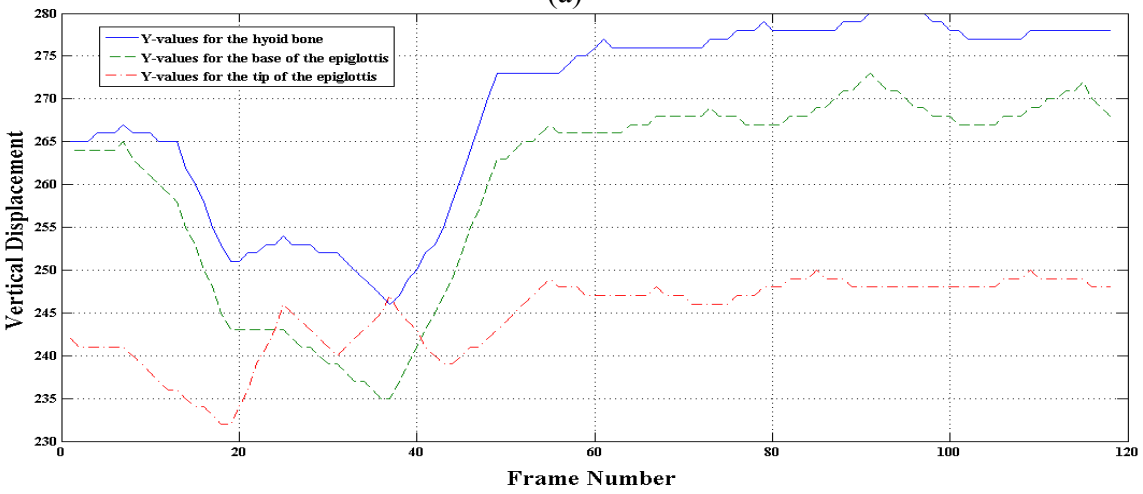


(b)

Figure A.1: Horizontal and Vertical displacements for the hyoid bone, base and tip of the epiglottis in patient 2

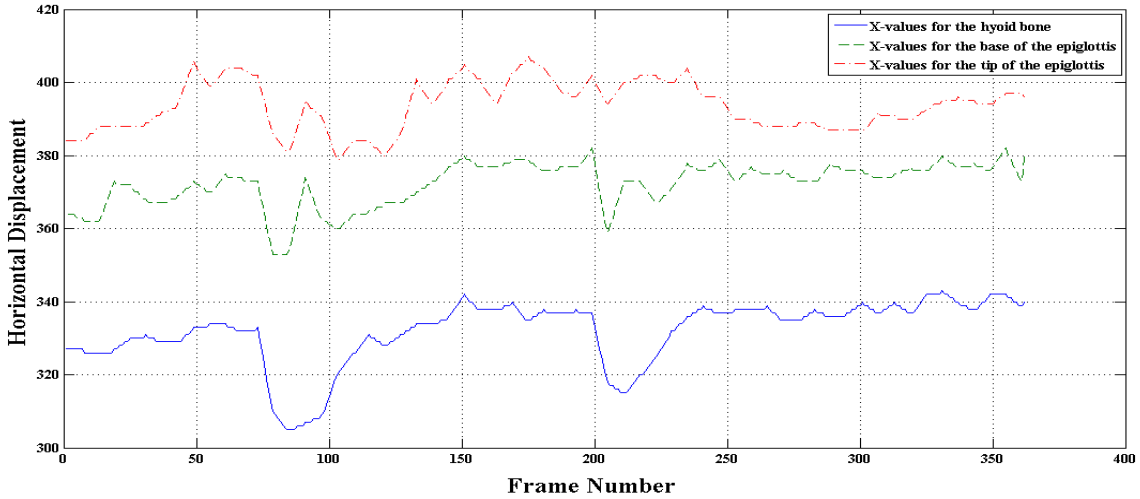


(a)

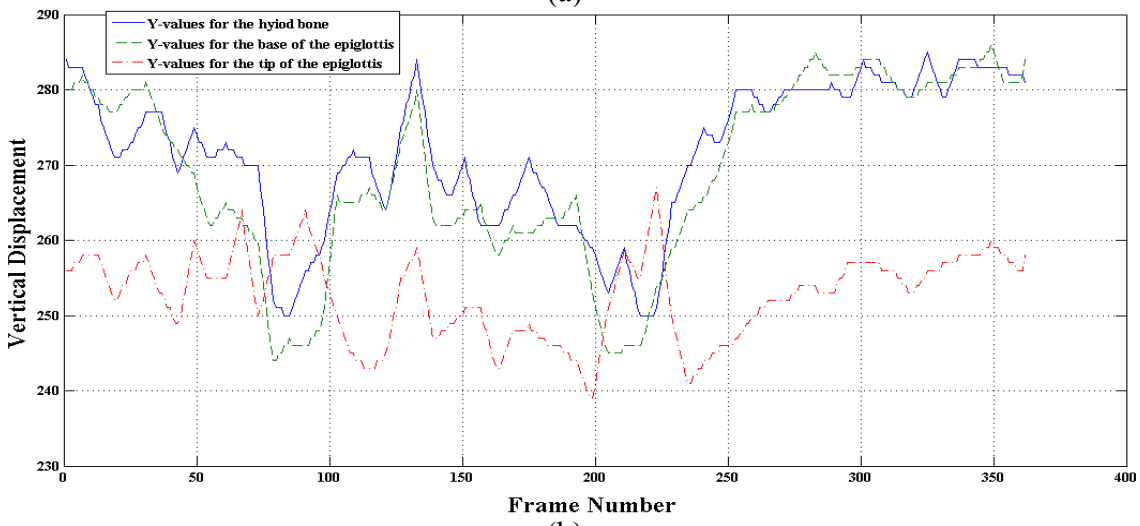


(b)

Figure A.2: Horizontal and Vertical displacement for the hyoid bone, base and tip of the epiglottis for patient 3

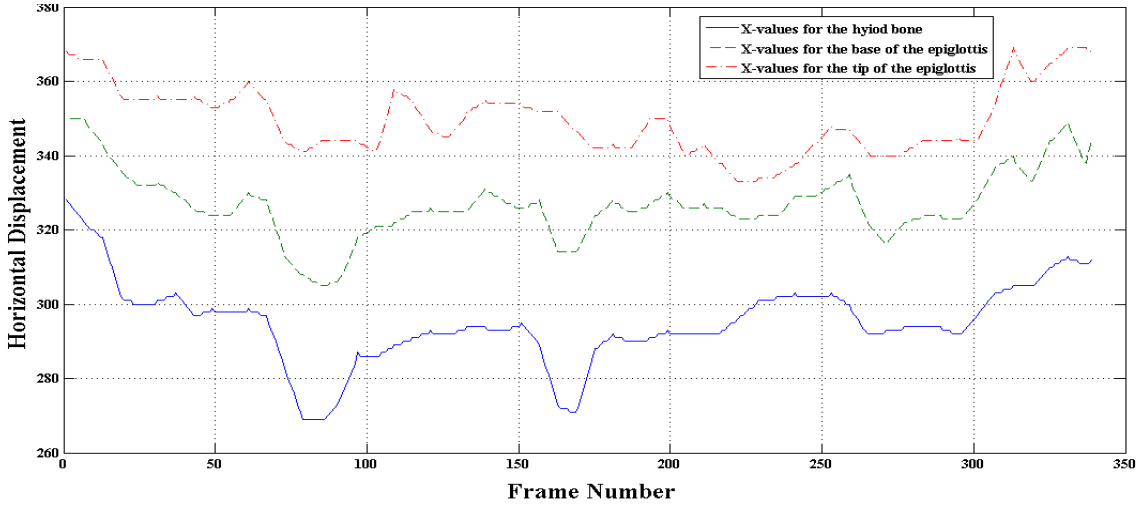


(a)

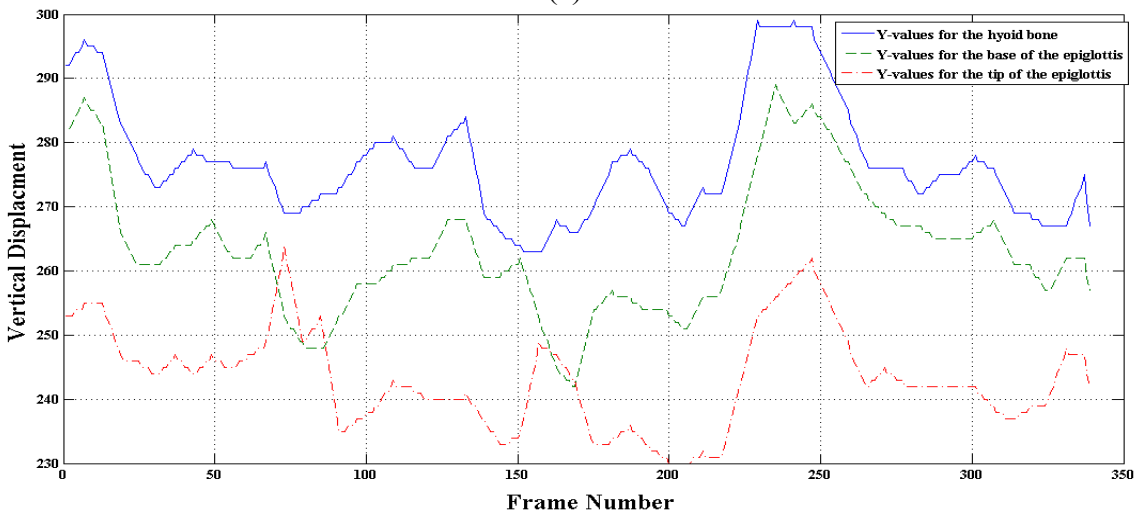


(b)

Figure A.3: Horizontal and Vertical displacement for the hyoid bone, base and tip of the epiglottis for patient 4

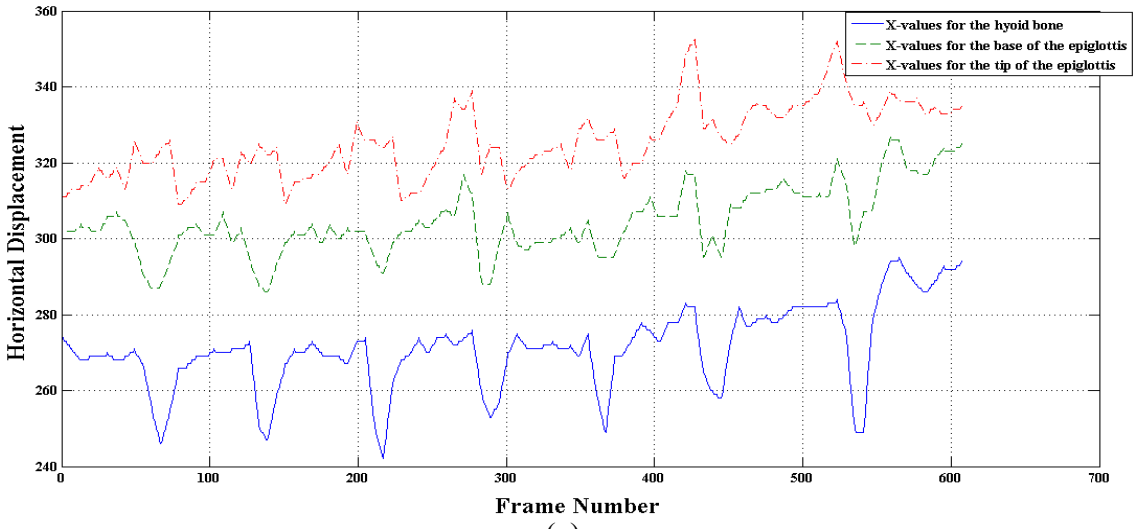


(a)

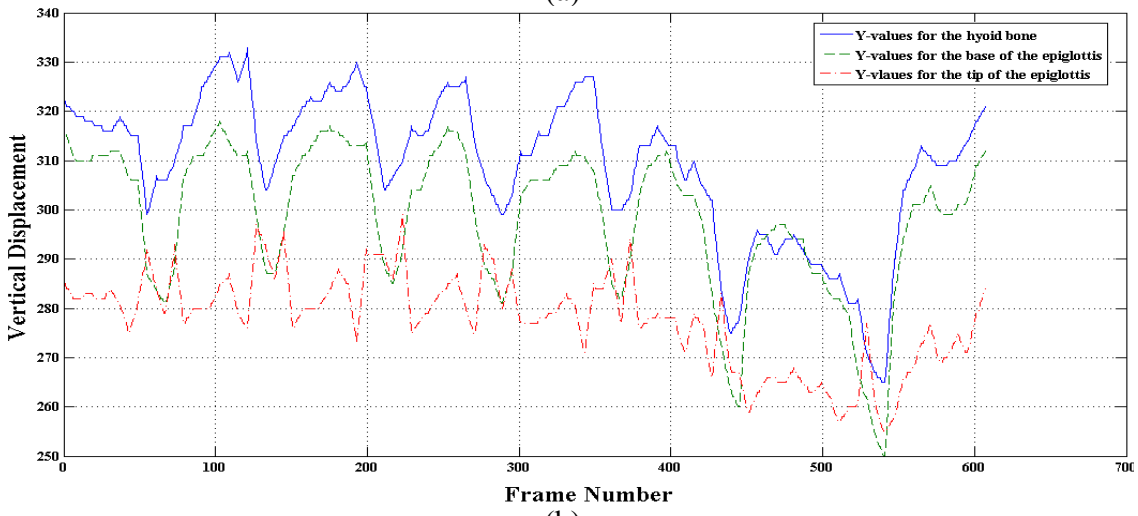


(b)

Figure A.4: Horizontal and Vertical displacement for hyoid bone, base and tip of the epiglottis for patient 5

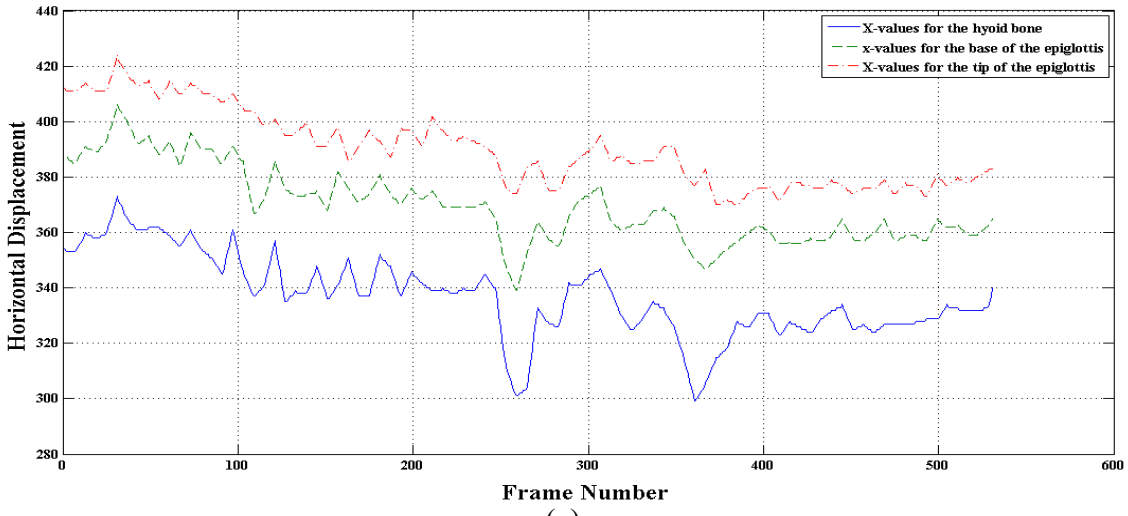


(a)

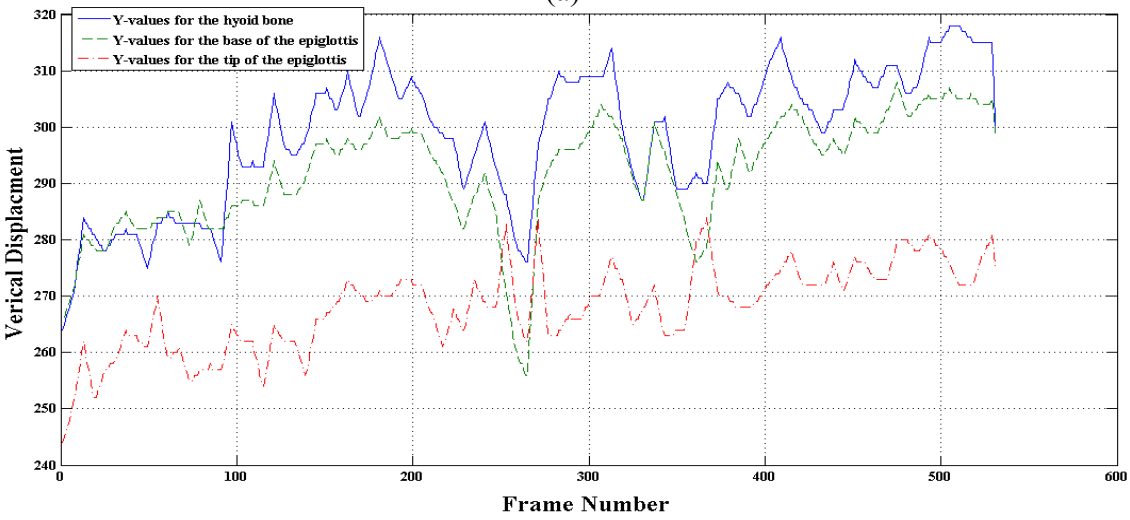


(b)

Figure A.5: Horizontal and Vertical displacement for the hyoid bone, base and tip of the epiglottis for patient 6

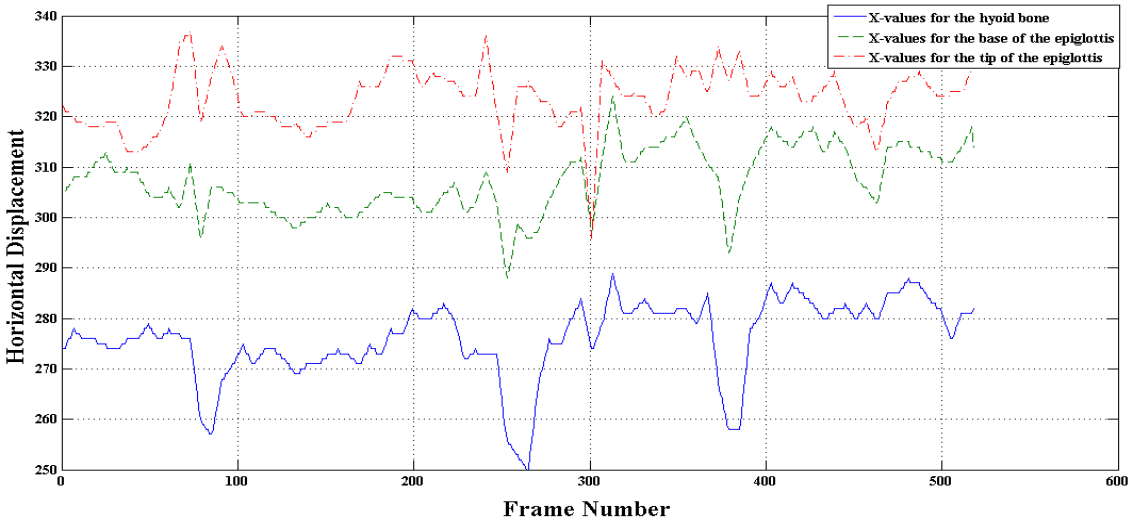


(a)

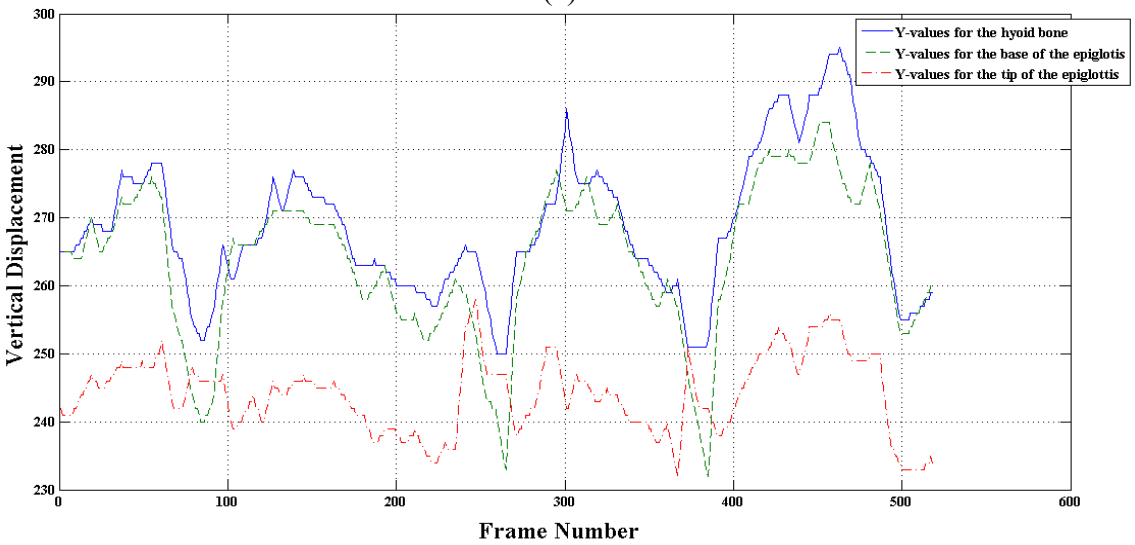


(b)

Figure A.6: Horizontal and Vertical displacement for the hyoid bone, base and tip of the epiglottis for patient 7

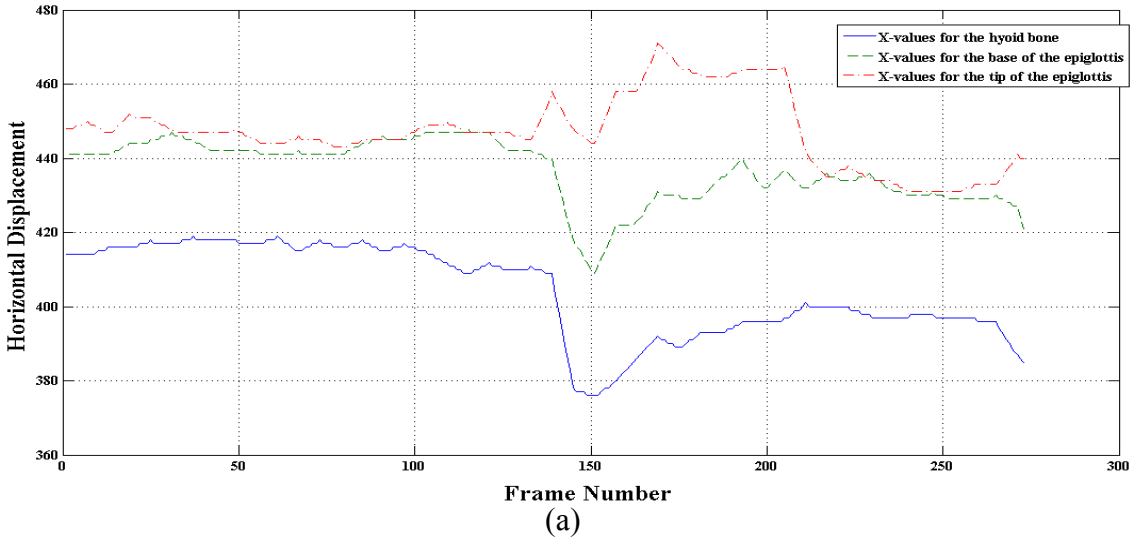


(a)

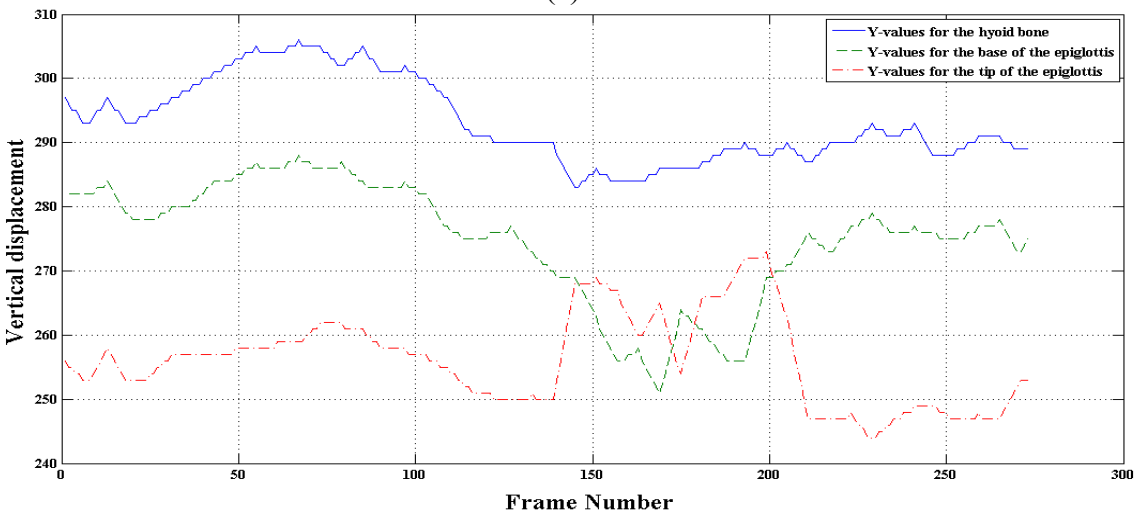


(b)

Figure A.7: Horizontal Vertical; displacement for the hyoid bone, base and tip of the epiglottis for patient 8

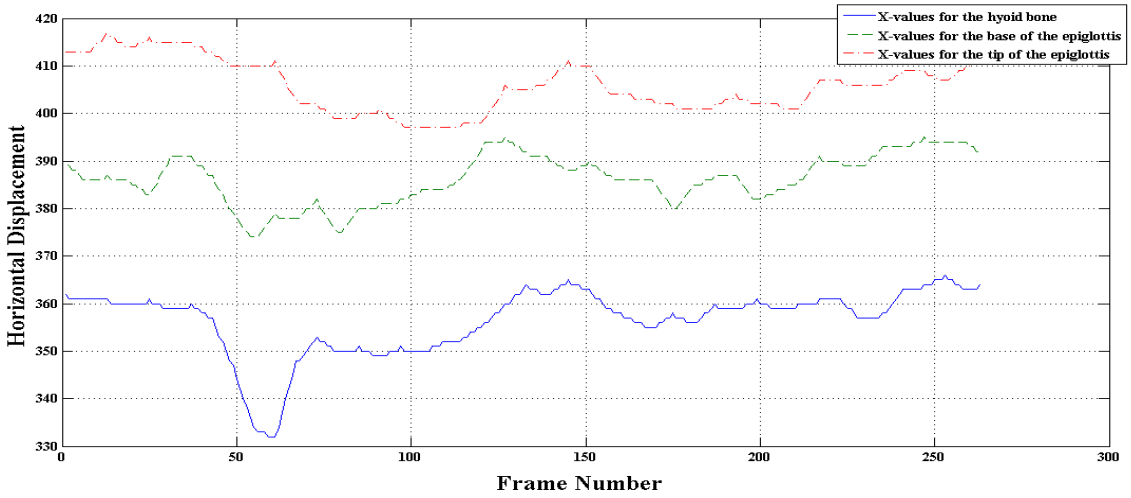


(a)



(b)

Figure A.8: Horizontal and Vertical displacement for the hyoid bone, base and tip of the epiglottis for patient 9



(a)

(b)

Figure A.9: Horizontal and Vertical displacement for the hyoid bone, base and tip of the epiglottis for patient 10

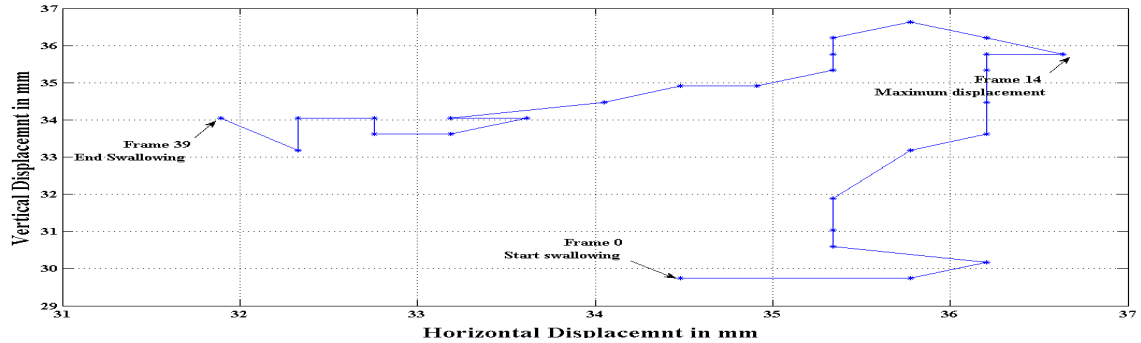
Appendix B. OpenCV Functions

Table B.1: OpenCV functions that are used and explanation

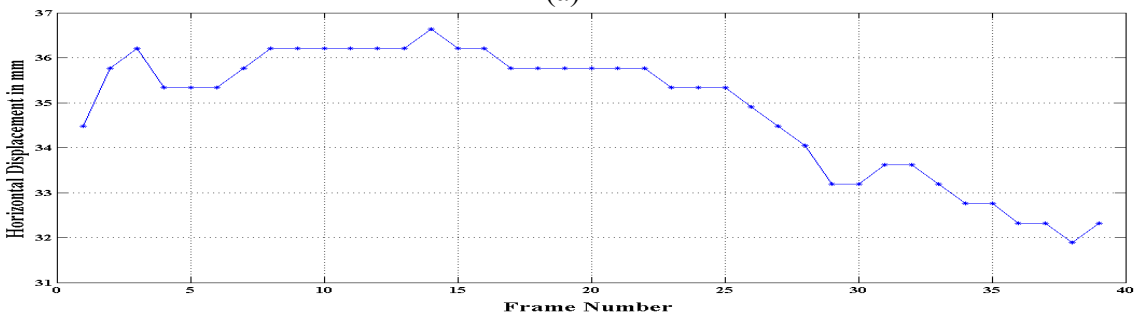
Function name	Explanation
cvCaptureFromFile	Initializes capturing a video from a file.
cvGetCaptureProperty	Gets video capturing properties.
cvCreateImage	Creates the header and allocates data
cvCopyImage	Copy data in the images
cvCloneImage	Makes a full copy of an image, including the header, data, and ROI.
cvSaveImage	Saves an image to a specified file.
cvReleaseImage	De-allocates the image header and the image data.
cvSetImageROI	Sets an image Region Of Interest (ROI) for a given rectangle.
cvResetImageROI	Resets the image ROI to include the entire image and releases the ROI structure.
cvLaplace	Calculates the Laplacian of an image.
cvRectangle	Draws a simple, thick, or filled rectangle.
cvCircle	Draws a circle.
cvLine	Draws a line segment connecting two points.
SurfFeatureDetector	Detects key-points and computes SURF descriptors for them.
SurfDescriptorExtractor	Computes the descriptors for a set of key-points detected in an image
BruteForceMatcher	Compute matching

drawMatches	This function draws matches of key-points from two images on output image.
cvCalcOpticalFlowPyrLK	Calculates the optical flow for a sparse feature set using the iterative Lucas-Kanade method with pyramids.
cvFitLine	Fits a line to a 2D or 3D point set.
cvEqualizeHist	Equalizes the histogram of a gray-scale image.
cvMatchTemplate	Compares a template against overlapped image regions.
cvGetQuadrangleSubPix	Retrieves the pixel quadrangle from an image with sub-pixel accuracy.
cvSetMouseCallback	Assigns callback for mouse events.
cvNamedwindow	Creates a window.
cvStartWindowThread	Allow OpenCV updating its windows automatically
cvShowImage	Displays the image in the specified window
cvDestroyAllWindows	Destroy all windows

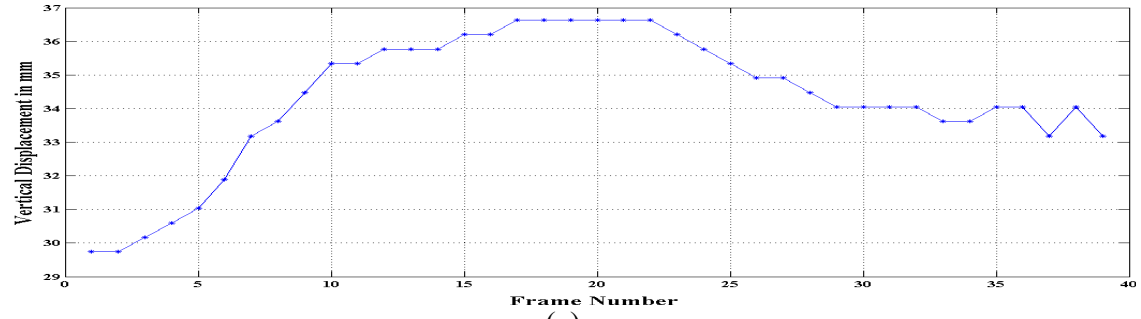
Appendix C. Extra Results For the Proposed Method



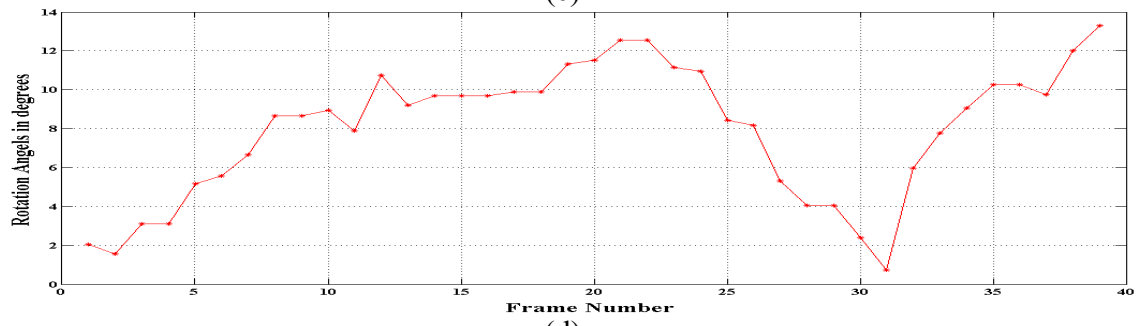
(a)



(b)

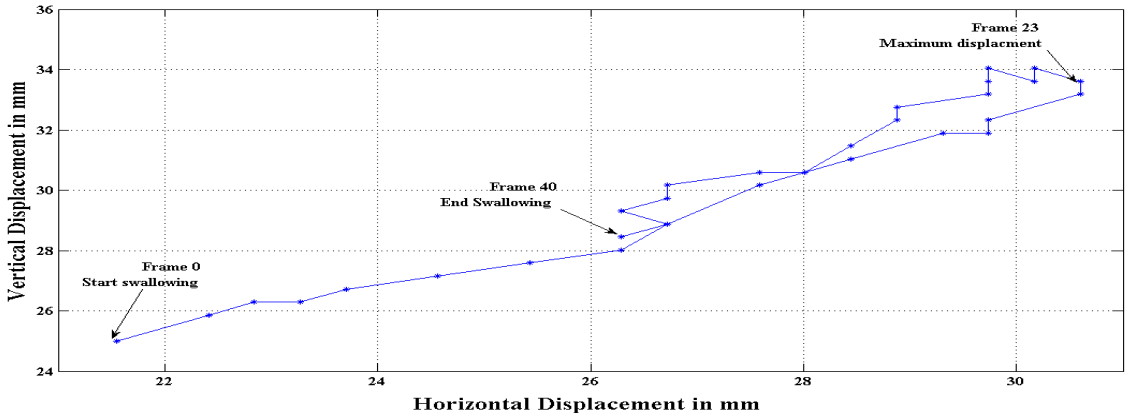


(c)

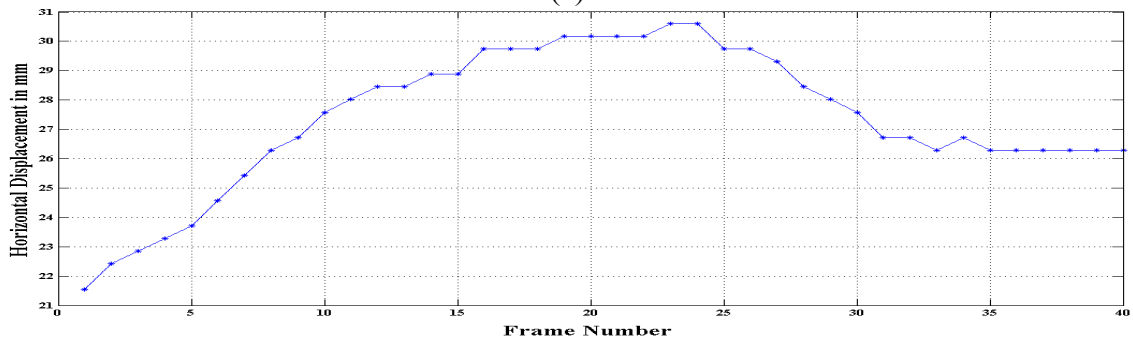


(d)

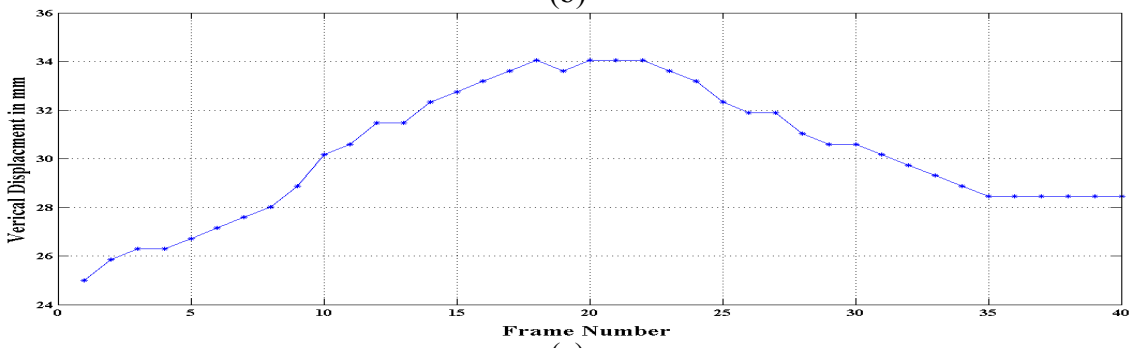
Figure C.1: First Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis



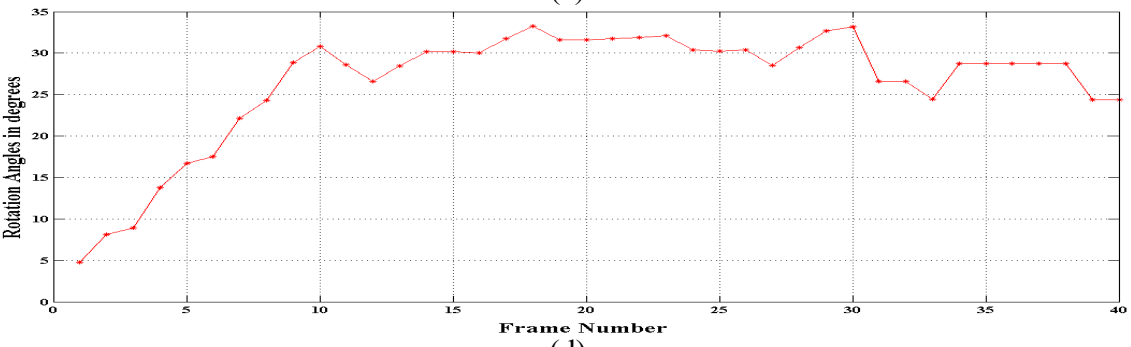
(a)



(b)

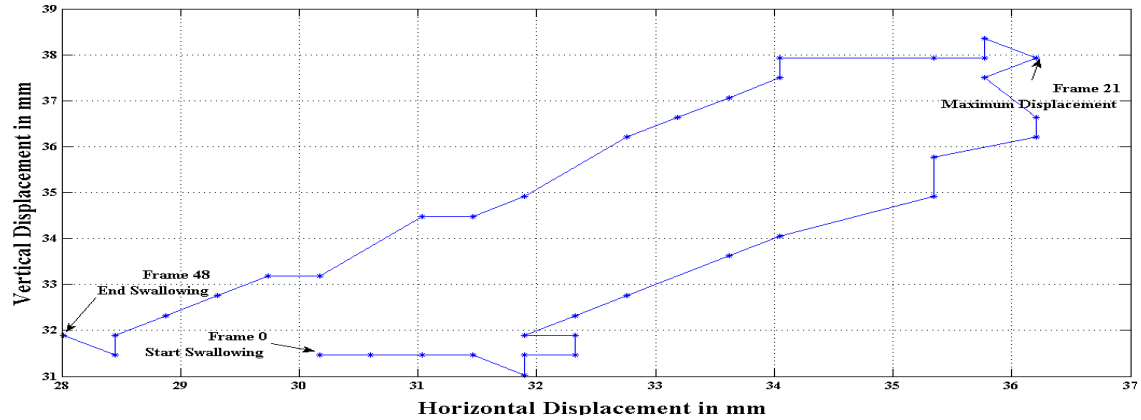


(c)

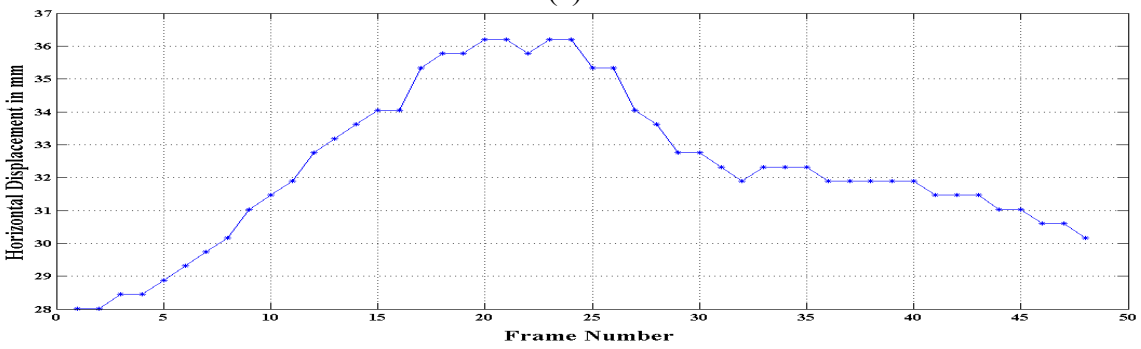


(d)

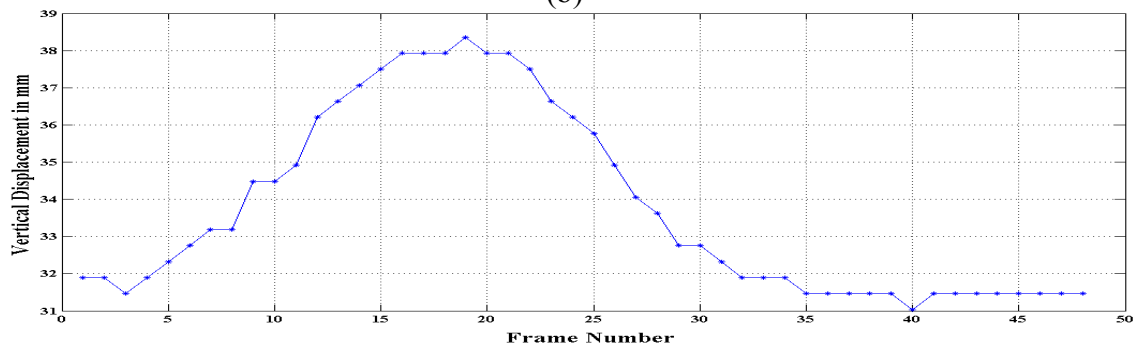
Figure C.2: Second Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis



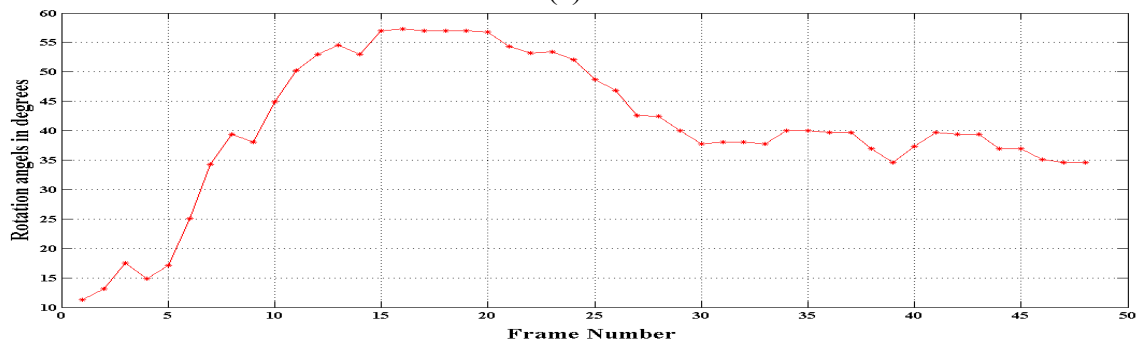
(a)



(b)



(c)



(d)

Figure C.3: Third Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis

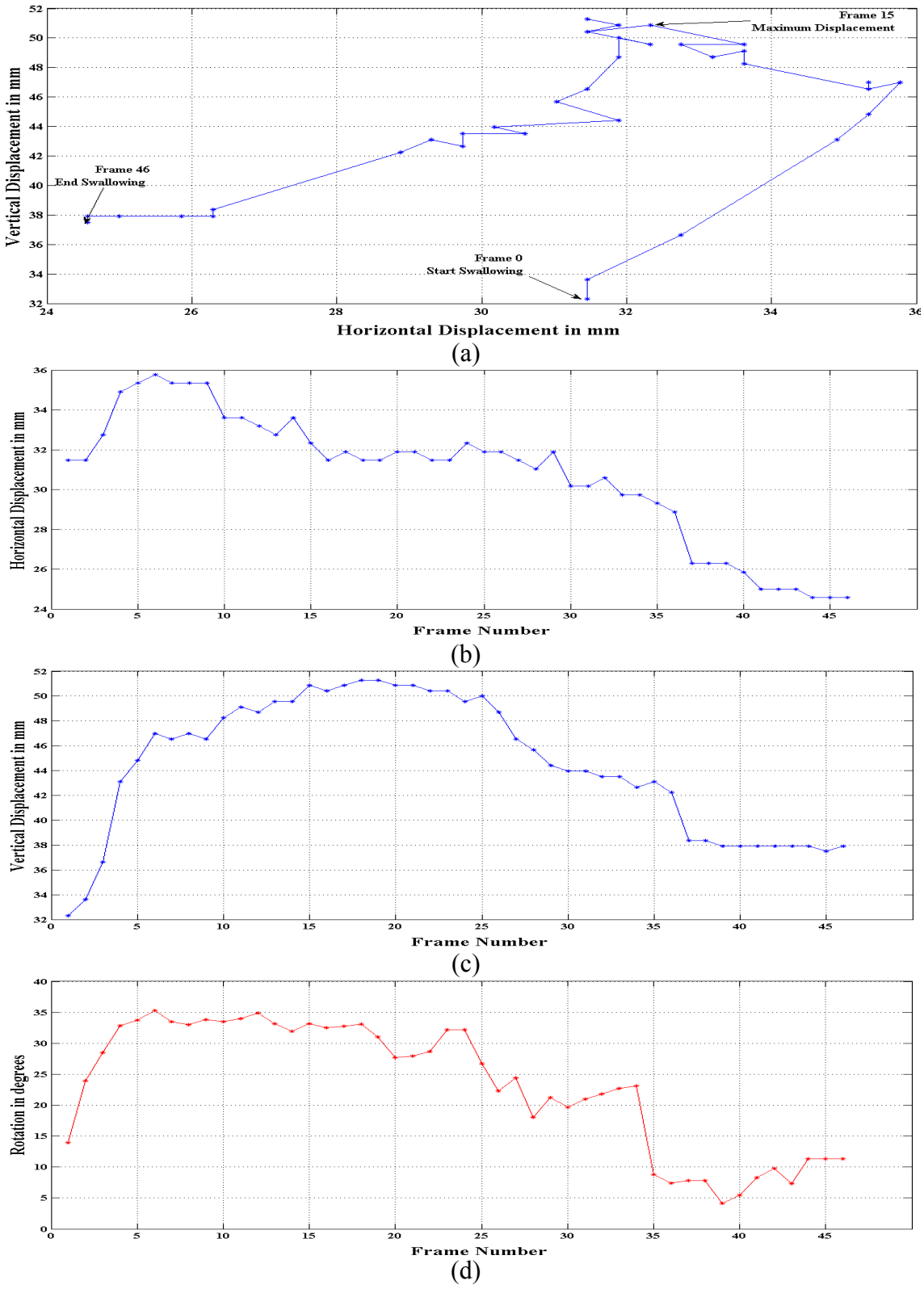


Figure C.4: Fourth Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis

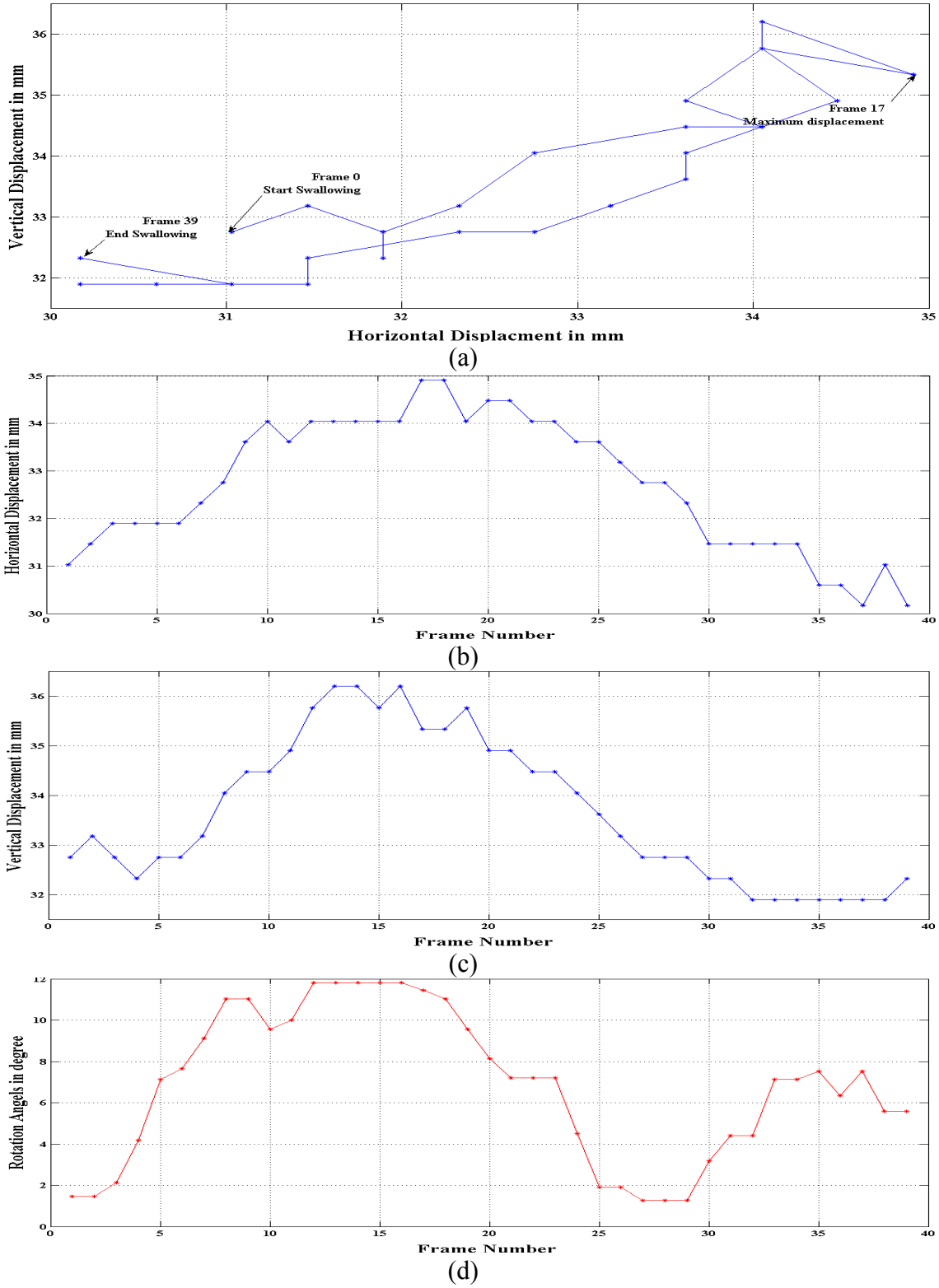


Figure C.5: Fifth Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis

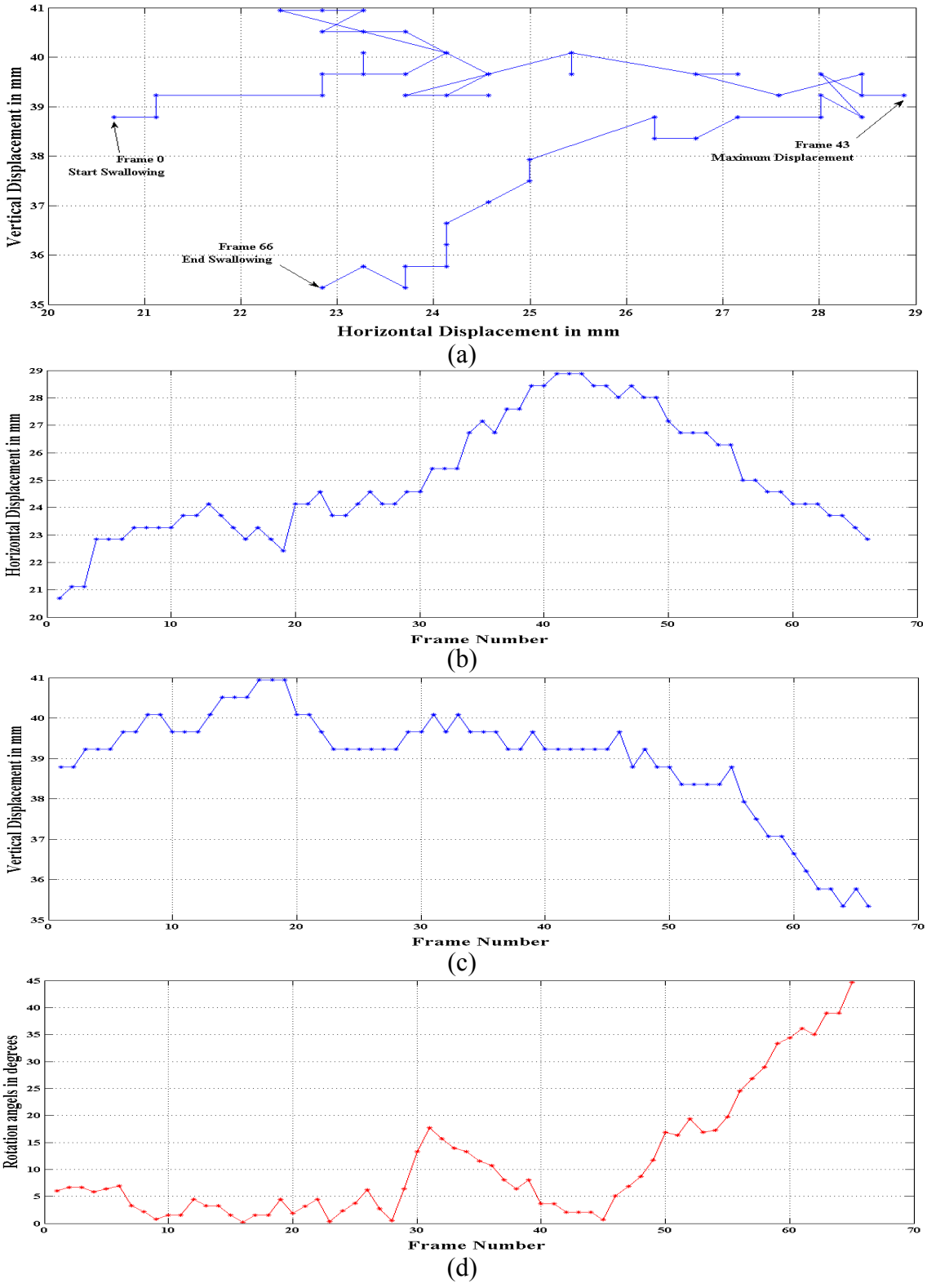


Figure C.6: Sixth Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis

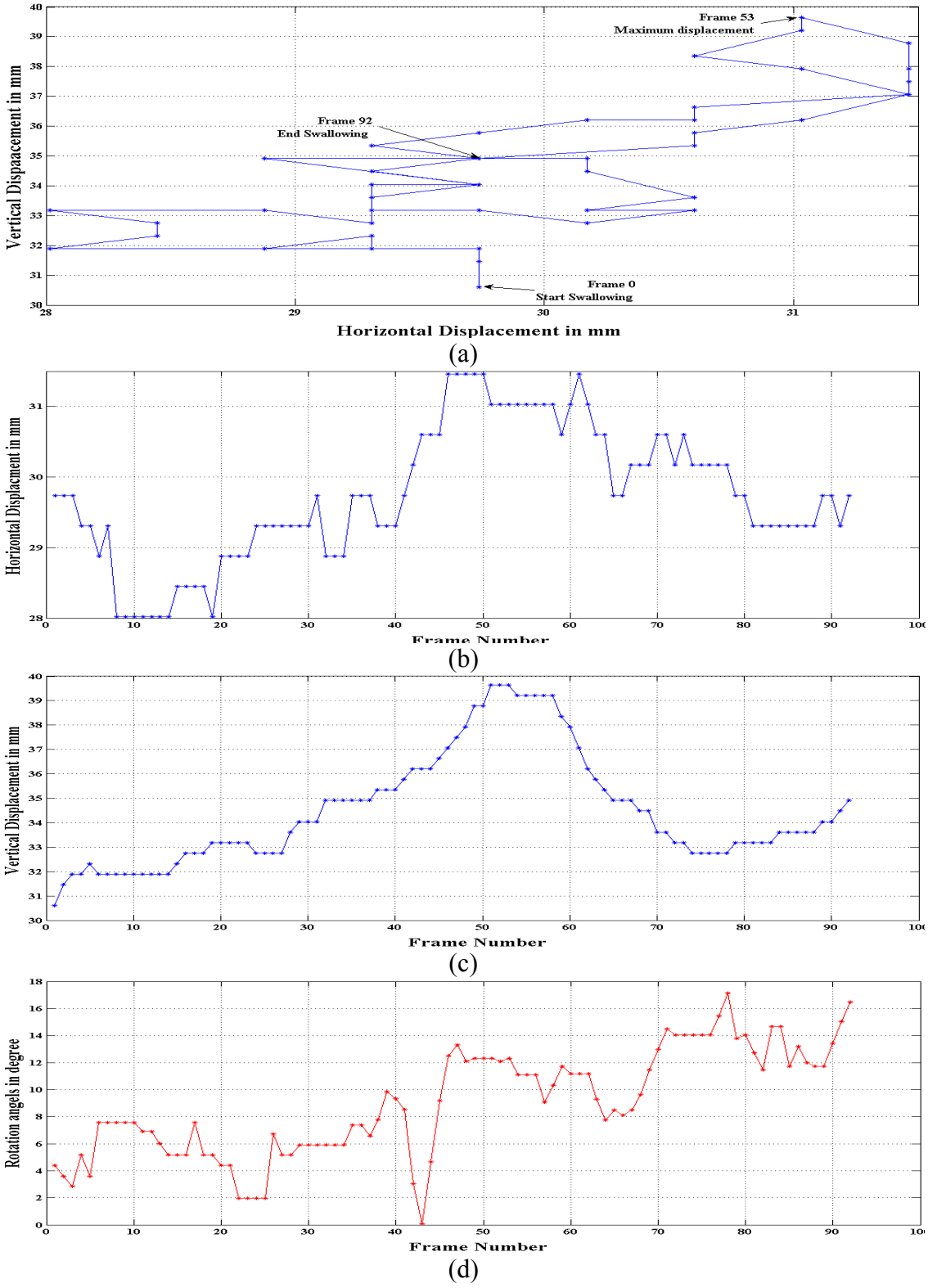
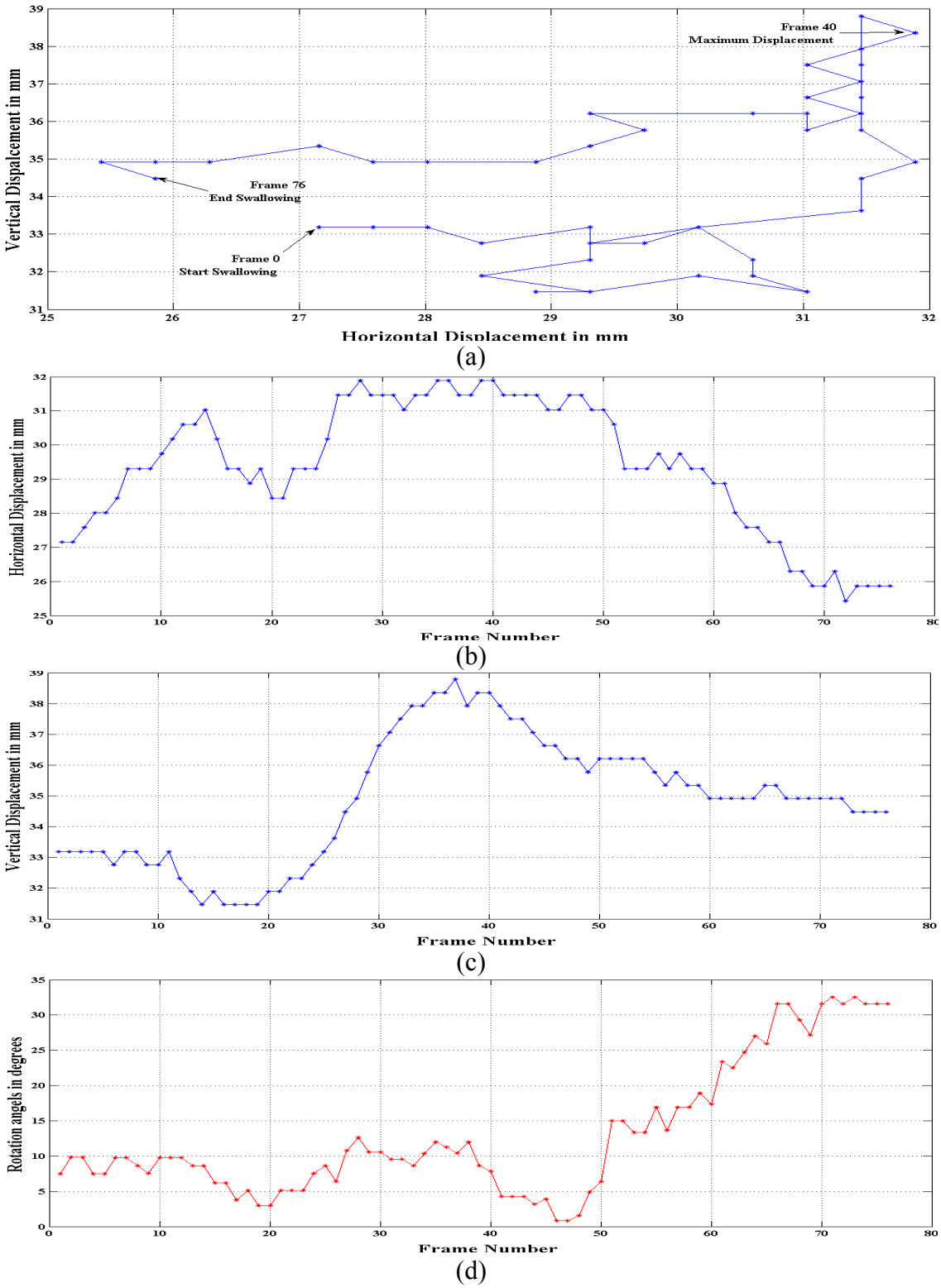


Figure C.7: Seventh Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis



FigureC.8: Eighth Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis

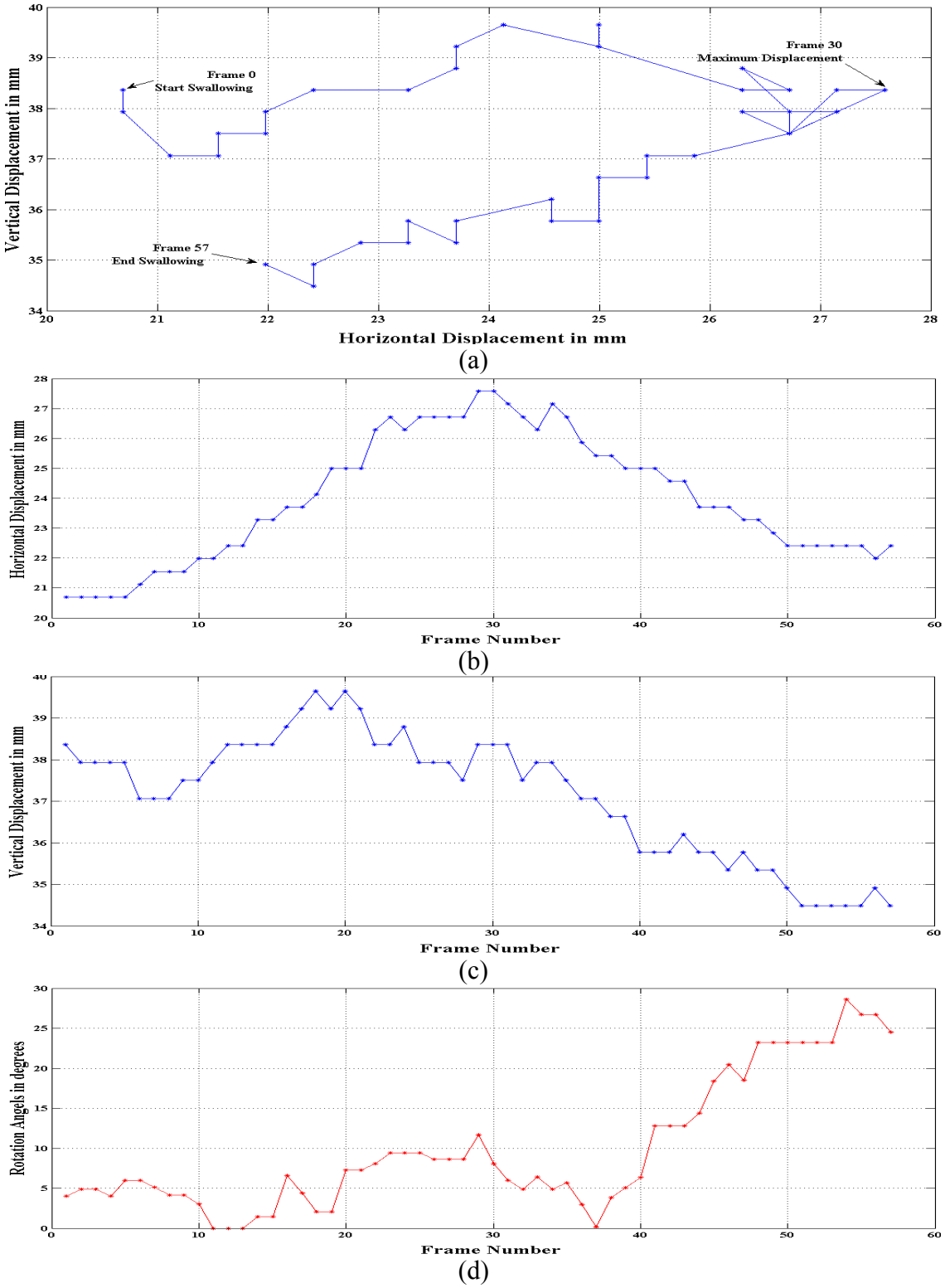


Figure C.9: Ninth Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis

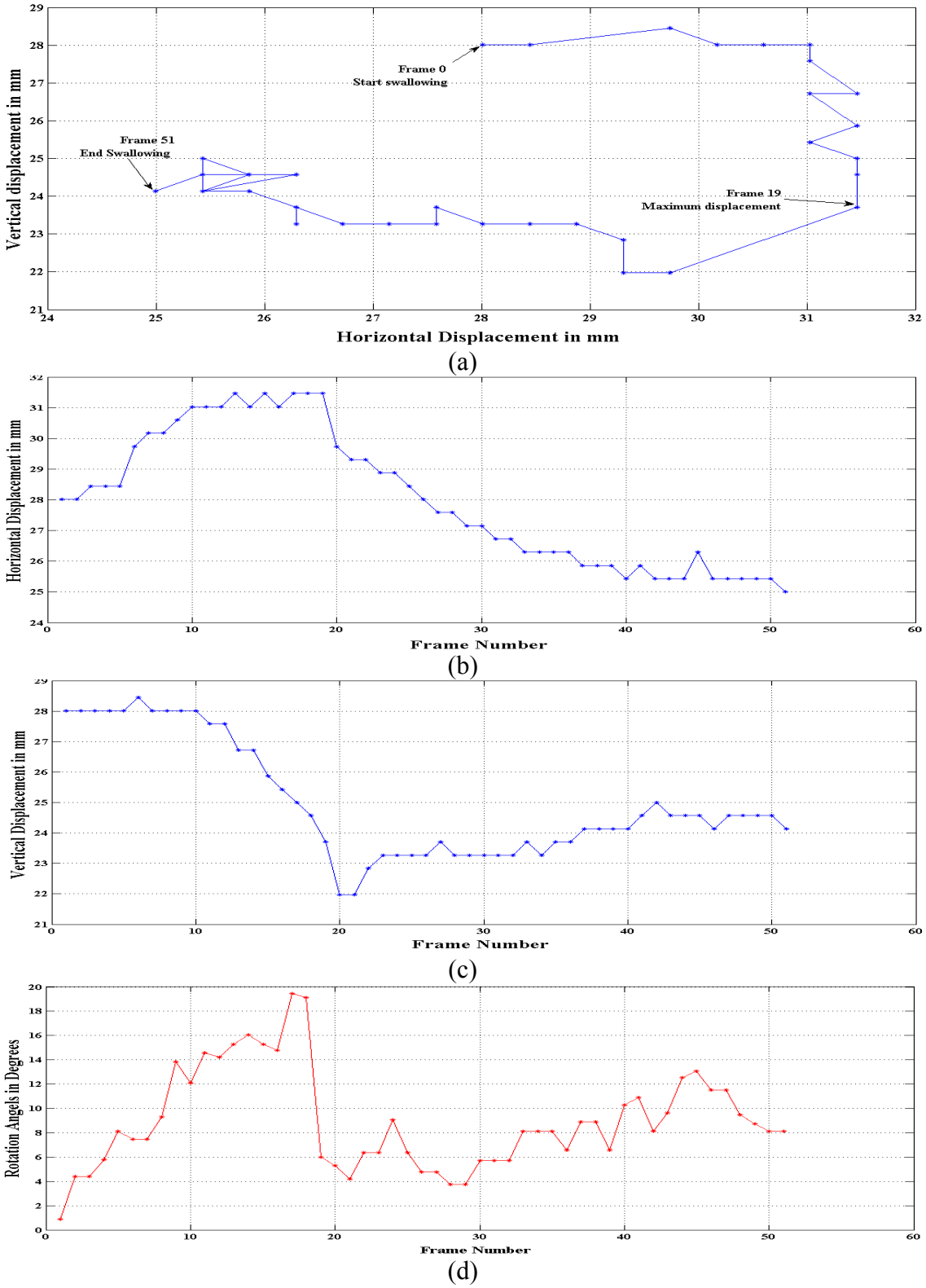


Figure C.10: Tenth Swallowing Process. (a) Trajectory of the hyoid bone. (b) Horizontal displacement. (c) Vertical displacement. (d) Rotation angels for the epiglottis

Curriculum Vitae

Name:	Seereen Noorwali
Place of Birth:	Makkah, Saudi Arabia
Post-secondary Education and Degrees:	Umm Al-Qura University (UQU) Faculty of Computers and Information systems Makkah, Saudi Arabia 2005 – 2009 B.Sc (Major Computer Sciences) CultureWorks ESL London, Canada 2010-2011
Related Work Experience	Teaching Assistant Umm Al-Qura University, Makkah 2009- 2010
Honour and Awards	Excellent and 1st Degree Honour in Computer Science Department Umm Al-Qura University, Makkah 2009 Scholarship for completing high education Umm Al-Qura University, Makkah 2010